

2 November 2007 | \$10

Science

AAAS ANNUAL MEETING
Science and
Technology
from a Global
Perspective
14–18 February 2008
Boston, Massachusetts

 AAAS



Donald Kennedy is the Editor-in-Chief of *Science*.

Approaching Science

SCIENCE COMMUNICATION IS A HOT TOPIC THESE DAYS. I KNOW OF TWO COMMITTEES exploring how science journalists and scientists can find new ways of working more effectively with one another. Programs such as the Aldo Leopold Leadership fellowships are educating young scientists about how to make their work more understandable to the public and to the mainstream media. These are common-sense responses to two important changes. First, the scientific disciplines are drilling deeper into the fine details of everything from atmospheric physics to the molecular basis of cell signaling. Acronyms and other shorthand indicia are used more often than ever, and even the titles of research papers are sometimes challenging. Second, science and technology are increasingly relevant to public policy, and unless those who speak for science can be understood, the policy decisions are likely to be wrong.

So how are we doing at *Science*? From readers' surveys, discussions with our Board of Reviewing Editors, and my mail, we get the same story. The language used in Reports and Research Articles is sufficiently technical and arcane that they are hard to understand, even for those in related disciplines. In conversations about this problem, I have heard the following: "It doesn't have to be like this; editors push hard to make authors compress their language too much." (Point taken; but allowing extra space for clarity in paper A makes less room available for the new findings in paper B, and pages also cost.) Or: "Can't you do a better job of teaching some of the scientists to write in a more accessible or understandable way?" (Another good point; but editors who do that sometimes find they edit for more clarity while authors insist that they are sacrificing precision and accuracy!)

It's clear that accessibility is a problem, because we're all laypeople these days: Each specialty has focused in to a point at which even the occupants of neighboring fields have trouble understanding each others' papers. We have already undertaken some measures that we think are helpful. Well before each issue, we gather News and Editorial staffs to consider which forthcoming papers should be accompanied by News or a Perspective, the latter of which provides a context for, or possible extensions of, new findings published in the same issue. We hope that this extra coverage improves comprehension.

But it's clear that to solve the underlying problem we need to do something more ambitious. Accordingly, we are initiating a new experiment, beginning on p. 766 in this issue. Each Research Article published this week and in the next five issues will be preceded by a one-page "Authors' Summary": an account, with one figure, of what the paper reports and what its conclusions are. Each author will have agreed to supply such an account and to let us improve its accessibility and clarity where needed. Our plan is for summaries of papers in physical science fields to be reviewed by our life-sciences editors and vice versa.

This experiment will affect both the print version of *Science* and *Science* Online. It is important for our readers to understand the purposes of the different ways in which we present research papers to improve understanding. The one-page summary is intended to make clear what the investigators did, how it was done, what the result was, and its significance. A Perspective, as always, will be complementary, putting the new result in context, exploring its impact, drawing in related information from other studies, and assessing how it changes the field.

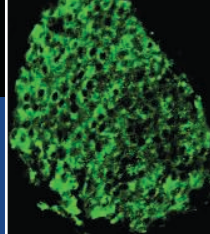
Obviously, we want to evaluate the outcome of this experiment, so we invite readers to look with special care at the first of these, in this issue, and at those that will follow for the next several weeks. Give us your opinion, either by returning a survey card, which will be in every issue that has summaries, or by going online to www.sciencemag.org/sciext/easurvey to fill out a short questionnaire. We look forward to hearing from you!

— Donald Kennedy



Mail or online survey:
[www.sciencemag.org/
sciext/easurvey](http://www.sciencemag.org/sciext/easurvey)

10.1126/science.1151603

Insight into
gestational
diabetes

729

A polymath
for the ages

733



A matter of words. Senator Barbara Boxer (D-CA) and CDC Director Julie Gerberding (right) confer before a hearing on health impacts of climate change.

CLIMATE CHANGE

CDC Director's Message on Risk Runs Afoul of White House Edits

Late on Monday evening, 22 October, the director of the U.S. Centers for Disease Control and Prevention (CDC) settled into a Washington hotel room for a ritual familiar to bureaucrats who regularly brief lawmakers: a late-night cram session with a three-ringed briefing binder. Julie Gerberding, was testifying the next morning before the Senate Environment and Public Works committee, which was holding the first-ever Senate session on the health impacts of climate change. Next to her was a 12-page summary of the scientific issues, prepared by agency researchers, to be distributed at the hearing.

But shortly before the hearing began, Gerberding learned that the White House Office of Management and Budget (OMB) had slashed the handout in half. Gone were details on how climate change could worsen allergic diseases, exacerbate deadly heat waves, and broaden the geographic range of infectious diseases. Gerberding didn't flinch. As she presented her 3-minute opening statement, a slide listed asthma, allergies, and harmful algal blooms as "potential negative health impacts." She expanded on those issues in response to questions from the committee chair, Senator Barbara Boxer (D-CA), and other members. The cuts didn't come up.

But the matter didn't end there. Within hours, reporters had their hands on the original version. It asserted, among other things, that the United States "is expected to see an increase in the severity, duration, and fre-

quency of extreme heat waves" and that studies showed higher urban temperatures could increase ground-level ozone, worsening asthma and "chronic lung diseases."

Why was her written testimony altered so severely? A White House spokesperson said the edits were done because details "didn't align with" the conclusions of the Intergovernmental Panel on Climate Change (IPCC), the group of thousands of scientists whose reports earned it this year's Nobel Peace Prize (*Science*, 19 October, p. 372). The president's science adviser, John Marburger, went into more detail 2 days later in a rare explanation of his role. Marburger said staff found "appropriate connections between climate change and human health" in the draft but criticized "several nuanced but important differences



Cutting-edge science. Science adviser John Marburger suggested removing parts of Gerberding's scheduled Senate testimony.

between the IPCC report's findings and the draft testimony."

Boxer called the editing the latest example of the White House "interfering [with] the scientific facts." And in private, CDC scientists were outraged. "They gutted it," said one. "For them to say that other agencies had a problem with it is just a weak cover ... for suppressing the science." Gerberding says she got across her message, namely, that "from a public health perspective that climate change is an important ... issue." But she told *Science* that she regrets that the incident has "raised questions about the credibility of the CDC science."

Central to the controversy is whether Gerberding had accurately represented what's known about the effects of climate change on global health. On heat waves, IPCC this year assigned "medium confidence" to the assertion that climate change has already led to "increased heat-wave related deaths." OMB removed words in the CDC handout describing the "likelihood of higher mortality" among the elderly from a heat wave. (In her oral remarks, Gerberding mentioned the more than 25,000 fatalities linked to the 2003 heat wave in Europe but focused on how U.S. officials were dealing with the problem.)

In terms of infectious diseases, the IPCC's 2007 report mentioned a recent study showing a *Vibrio* bacteria outbreak in oysters "linked to atypically high temperatures in Alaskan coastal waters." It noted that extreme spring rains in North America and Europe "may explain" patterns in cryptosporidiosis bacterial disease. Both examples were cut from the handout and went unmentioned at the hearing.

Gerberding, an infectious-disease scientist, says her role didn't require her "to go into detail on every single complication of climate change. ... Rarely, if ever, does the written testimony get read [by the public]." But Rick Piltz, formerly of the White House Climate Change Science Program, disagrees. "The official statement on a problem from the top-ranking public health official to Congress? It matters."

Those familiar with the process say the testimony, submitted to the White House 5 days before the hearing, was vetted by agencies including NASA, the Environmental

CREDITS (TOP TO BOTTOM): ALEX WONG/GETTY IMAGES; BRENDAN SMIALOWSKI/BLOOMBERG NEWS/LANDOV

Downloaded from www.sciencemag.org on November 4, 2007



Protection Agency, and Marburger's Office of Science and Technology Policy (OSTP). But Gerberding said none of the comments she saw were "showstoppers." OSTP, for example, took issue with specific wording on hurricanes, food scarcity, and mental health. An OSTP official says that the office "had no problem" with sections on infectious diseases and heat waves but that "I don't know what

happened to our edits," which were submitted on Monday. Marburger's statement said that OSTP "did not seek to redact sections."

Gerberding says she's skeptical of the OSTP assertion that OMB's heavy editing was necessary because of insufficient time to reconcile the various comments, calling that "just a hypothesis." She's asked her staff to review the 12-page draft and the vetting process "to

make sure this doesn't happen again."

If the goal was to suppress the issue, extensive media coverage suggests that the attempt failed. Gerberding says she'd be pleased if the flap about her testimony leads to more talk about CDC's role in studying climate change, calling it "a silver lining." She says she has tried to "get this issue on the map for public health, and I will continue to do so." —ELI KINTISCH

NATURAL DISASTERS

Drone Spy Plane Helps Fight California Fires

SAN DIEGO, CALIFORNIA—One unsung hero in the weeklong battle against the massive wildfires that devastated southern California last week was an unmanned Predator B airplane, originally designed to gather intelligence for the military. By providing firefighters with up-to-the-minute data on the many conflagrations, it helped in the coordination of firefighting efforts that spared many lives and structures—including the venerable Palomar Observatory.

The plane now flies for NASA and the U.S. Forest Service, which have outfitted it with state-of-the-art, thermal-infrared sensors to help firefighters decide where to do battle. The 5-year demonstration project, dubbed Ikhana after the Native American Choctaw word for "aware" or "intelligent," transmitted hot-spot data to a satellite. Within 15 minutes, fire command centers all over San Diego received color-coded Google Map images that indicated the temperature on the ground at different locales. "The fire commanders love the data," says aerospace engineer Brent Cobleigh, who leads the project from NASA's Dryden Flight Research Center inside Edwards Air Force Base.

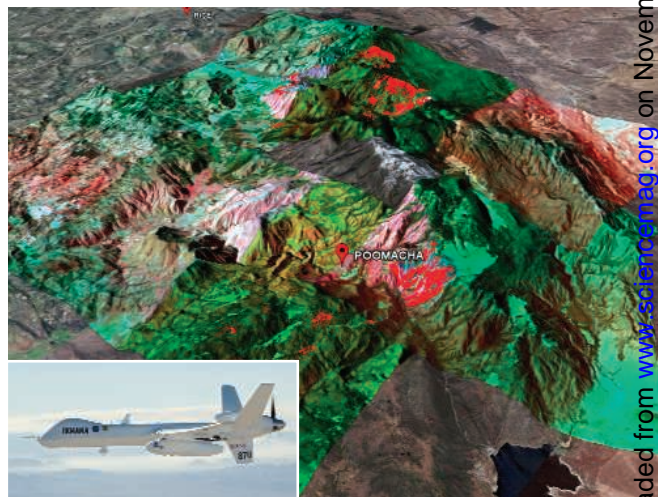
The Forest Service for 40 years has relied on images from piloted planes, notes Everett Hinkley, a lead investigator on the project who is based at the Forest Service's Remote Sensing Applications Center in Salt Lake City, Utah. Ikhana represents a technological step up, he says, with a novel sensor that can penetrate smoke and delineate hot spots by detecting temperature differences as small as 0.5° up to 1000°C. "We can get very finite and fine temperature discrimination from surface events," says Vincent Ambrosia, Ikhana's principal investigator at NASA's Ames Research

Center at Moffett Field, which developed the sensor.

Ikhana's near-real-time information is also a huge improvement. Piloted airplanes don't have this equipment and can only transmit their data by airdrops or after they land, so several hours often pass before firefighters receive it. "In an environment like southern California, with very high winds and fires moving rapidly, thermal imagery can be out-of-date quickly," says Hinkley. NASA hopes eventually to transfer that capability to piloted planes.

Ikhana's pilots sit in a trailer at NASA's Dryden Center, so the plane can fly into dense smoke plumes or high winds without putting a crew's life in danger. Ikhana can circle a hot spot for hours on end, for example, and fly for as many as 30 hours straight, assignments that would severely stress an onboard crew. Ikhana made its first flights over fires in August, but the dozen fires that burned in San Diego County last week put it to a tougher test. "Everything's happening a lot faster," said Ambrosia on Friday, 26 October, the third day of Ikhana's San Diego mission.

That day, the plane flew over the Poomacha fire, which came within several kilometers of the Palomar Observatory. Owned and operated by the California Institute of Technology, the observatory's five telescopes include the 508-cm "Big Eye." Astronomer W. Scott Kadel, a spokesperson for the observatory who had to evacuate his own home nearby, said, "I don't know of a fire that's ever been this big of a threat here."



Getting the picture. Thermal-infrared images from the unmanned Ikhana (*inset*) helped firefighters battle the Poomacha fire near San Diego's Palomar Observatory.

Fortunately for the observatory, the Poomacha fire was one of the last to break out. "Many other fires had reached their conclusion, and support from across the state had arrived," says Kadel. "If we had been first to break out, we'd be talking about what a great place this used to be."

With the demonstration project coming to an end, researchers are awaiting word on a funding request to sustain the effort. Money isn't the only hurdle, however. Strict rules about flying unmanned aircraft in prime national airspace (Ikhana flies at an altitude of 7 km) make it difficult to secure permission for each flight. And NASA's mission to develop these technologies doesn't cover responding to natural disasters. That puts Ikhana's future as a firefighter up in the air. —JON COHEN

MEDICINE

Tumor Suppressor May Also Affect Gestational Diabetes

Stanford University's Seung Kim was treating cancer patients when he first encountered a clue about the biology of diabetes. The patients suffered from a rare inherited syndrome called MEN1, or multiple endocrine neoplasia type 1, which triggers tumors in endocrine glands. In some cases, beta cells in the pancreas multiply and overproduce insulin.

This reminded Kim of a very different state: pregnancy. There, placental signals and the burden of a fetus lead to insulin resistance and a demand for more hormone. An unlucky 4% of women cannot produce enough and are diagnosed with gestational diabetes.

On page 806, Kim, a developmental biologist, and his Stanford colleagues describe a connection in mice between *menin*, the protein encoded by the *MEN1* gene, and gestational diabetes. *Menin* had long been known as a tumor suppressor; MEN1 occurs when the protein is lacking. Now, Kim and his colleagues report that in healthy animals, *menin* inhibits the replication of beta cells. Dialing it down allows beta cells to proliferate, increasing insulin supply. It's a clever insight, says Robert Sorenson of the University of Minnesota, Minneapolis, who studies how these cells adapt to pregnancy: "I knew about multiple endocrine neoplasia, for years and years, ... but I never thought about linking it to pregnancy." If its role holds up in humans, *menin* may offer clues about gestational diabetes and the ways beta cells regenerate.

Kim's group first examined what happens to *menin* during pregnancy. In normal mice, *menin* levels drop and climb back up after the animal gives birth. To test whether this was driving beta cell expansion, Kim and his colleagues created transgenic mice whose *menin* production didn't decline during pregnancy. When these mice became pregnant, their glucose rose to levels comparable to those in gestational diabetes. Autopsies confirmed that when *menin* didn't fade in pregnancy, beta cell proliferation was impaired, and the ani-

mals didn't produce the insulin they needed.

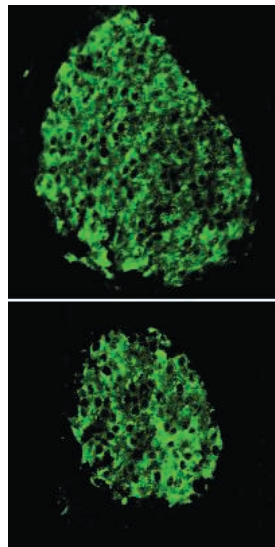
In mouse and human cells, Kim's team found that prolactin, a hormone that rises during pregnancy and that Sorenson had found triggers beta cell proliferation in rodents, was what caused *menin* levels to fall. This, says Sorenson, suggests a distinct chain of events: A boost in prolactin causes a drop in *menin*, which increases beta cell mass. The discovery is "quite different than the tumor story," because here the normal functions of *menin*

take center stage, says Stephen Marx of the National Institute of Diabetes and Digestive and Kidney Diseases in Bethesda, Maryland, who helped lead the team that cloned the *MEN1* gene. Still, because the work is in mice, "it's hard to extrapolate too much," Kim cautions.

The work suggests that *menin* might become a tool for developing beta cells in diabetes patients who lack them. "Up until about 2004, everybody said that you can't get new beta cells from replication; if you lose your beta cells, too bad," says Andrew Stewart, an endocrinologist at the University of Pittsburgh in Pennsylvania. The find that suppressing *menin* induces beta cell expansion is "one small step in a long quest" to understand beta cell replication, says Yuval Dor, a

developmental biologist at the Hebrew University-Hadassah Medical School in Jerusalem, Israel. Admittedly, devising a *menin*-based therapy could be complex, because it might interfere with *menin*'s role as a tumor suppressor.

Menin's relevance to gestational diabetes remains unclear. Some studies suggest that women who are diagnosed with gestational diabetes may have had insulin deficits before pregnancy. "You happen to intercept them" once they're pregnant, says Thomas Buchanan, an endocrinologist at the University of Southern California in Los Angeles, who doubts that *menin* is "an important contributor" in human gestational diabetes. Still, that's something scientists won't know until they assess *menin* in people.



Supply problem. Insulin-making beta cells in normal pregnant mice expand to meet demand (top), but not in mice with higher levels of *menin* protein (bottom).

New Scrutiny on Vaccine Trial

Stunned by the decision to halt a large trial of a promising AIDS vaccine made by Merck & Co. in September (*Science*, 5 October, p. 28), researchers are now confronting a more shocking prospect still: Vaccinated people somehow may have become *more* susceptible to HIV infection than participants who received a placebo. Experts caution that no new data have emerged, simply a new interpretation of the same findings.

An independent Data and Safety Monitoring Board (DSMB), required for certain federal clinical trials, recommended stopping the 3000-person trial, under way in North and South America, the Caribbean, and Australia, after an interim analysis of half the participants revealed that 19 people who had received at least two doses of the vaccine became infected versus 11 people in the placebo group. Because the vaccine had failed, a trial of the same preparation in South Africa was also stopped, and the DSMB overseeing that study suggested that researchers notify vaccinated people of the potential increased risk they now face, although the difference was not statistically significant. "We had two DSMBs that looked at the data differently," says Lawrence Corey, a study leader with the University of Washington, Seattle. Corey says there were "geopolitical considerations" that led the South African DSMB to "err on the side of safety." Study organizers will meet in Seattle next week to closely review the data.

—JON COHEN

French GM Crops Halted

PARIS—French President Nicolas Sarkozy delighted environmental activists last week by announcing a moratorium on the commercial cultivation of genetically modified (GM) crops. The temporary ban was one in a series of sweeping environmental policies announced at the end of a 2-day roundtable meeting on the environment here. Sarkozy said he wants to suspend cultivation of Monsanto 810, an insect-resistant maize and the only GM crop currently grown in France, until a new "High Authority" has been set up to weigh its risks and advantages. EuropaBio, an industry lobby group, said the plan violates European rules; a spokesperson for the European Commission confirms that countries cannot ban the European Union-approved strain but says that the commission is awaiting details. André Choulika, a board member of industry group France Biotech, says the authority could help inject more science into the GM debate, and if it rules on 810 before the next planting season, the ban's impact could be minimal.

—MARTIN ENSERINK

U.S. DEFENSE BUDGET

Pentagon Research Chief Seeks 11% Boost in 2009

U.S. academic lobbyists have long urged the Department of Defense (DOD) to boost its investment in science and technology (S&T) to enhance the nation's strategic and economic competitiveness. But although the agency's overall S&T portfolio has grown modestly since 2002, funding for basic research has stagnated. Now, in an unprecedented step, the Pentagon's research boss is seeking an 11% increase in DOD's S&T budget for the next fiscal year. Observers say the request, which Pentagon officials appear eager to publicize, improves the long-term prospects for more defense-related research despite the continued cost of the wars in Iraq and Afghanistan.

The proposal comes from John Young, director of Defense Research and Engineering (DRE). In a 24 August memo to Defense Secretary Robert Gates, which recently became public, Young warns that the agency's S&T investment "may be inadequate to meet the imposing security threats that challenge the nation." Citing the modernization of China's military, the decline in basic research by corporations, and the increasing sophistication of terrorist organizations, Young proposes establishing new S&T initiatives worth \$1.2 billion in the 2009 fiscal year (FY). Congress has yet to

complete action on this year's request for \$10.8 billion for the program.

At the top of Young's list is a \$300 million investment in foundational sciences, covering areas such as biosensors, photonic crys-

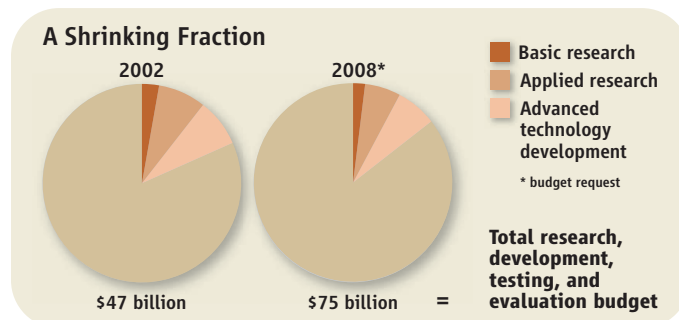
research spending (*Science*, 21 October 2005, p. 423). Adds Alan Shaffer, one of Young's senior executives: "We're not getting the same type of pushback we used to get. Yes, there are a lot of claimants for dol-

lars. But there's a groundswell of support for the idea."

Nonetheless, everybody admits that the proposed increase faces a big hurdle. "The commitments the department has made extend far beyond available resources," says Loren Thompson of the Lexington Institute, a Washington, D.C.-based nonprofit. For example, he says, the Air Force urgently needs 321 F-21 fighter aircraft, but "there's not enough money in the 2009 proposed budget to purchase even half" that number. "What

the services want right now is hardware, not science," he says.

A congressional staffer agrees that a big research boost next year is unlikely. "Just because Mr. Gates asked for the proposal and Mr. Young produced it does not make it a slam dunk," he says. But he believes that the plan prepares the ground for bigger increases down the road. Lobbyists also note that Young is awaiting congressional approval to become the agency's under secretary for acquisition, technology, and logistics, which would give him oversight of a \$157 billion budget that includes DRE. **—YUDHIJIT BHATTACHARJEE**



Smaller sliver. Spending on science and technology hasn't grown as fast as the rest of the military's research, development, testing, and evaluation budget.

tals, and the computing sciences. That represents a 20% hike in the agency's \$1.5 billion basic research budget. "We are pleased to see this discussion occurring within the higher levels of the Pentagon," says M. Matthew Owens of the Association of American Universities, adding that the DRE director hasn't asked for such a hike in recent history. "We are optimistic that it will have some impact on the FY '09 budget."

Supporters point to the fact that the memo was requested by Gates, a member of a National Academies panel that in 2005 called for a 10% increase in basic defense

CLIMATE CHANGE

California Stirs a \$600 Million Pot of Solutions

California researchers could soon be able to tap a 10-year, \$600 million climate initiative. The project would create the California Institute for Climate Solutions to foster research so the state can meet strict greenhouse gas emissions limits enacted over the past 2 years. The president of the state's public utilities commission (PUC), Michael Peevey, recently announced that PUC is looking at the proposed institute as a way to help meet the new targets. The commission is weighing a plan to finance it through a \$1-a-month hike in electricity rates.

"This is really exciting to see," says Daniel Kammen, an energy policy expert at the University of California (UC), Berkeley, who views the institute as a way to translate climate goals to action. "It will really put financial

muscle behind the climate-change laws." Omar Yaghi, a chemist at UC Los Angeles, who works on materials capable of separating carbon dioxide from power-plant emissions, likes the idea that utility commissioners are paying attention to more than just the industry's bottom line. "I'm really happy to hear the PUC is taking the initiative on this," Yaghi says.

The institute's design is still in flux. But Kammen and others say it's likely to focus on a range of projects that offer near-term energy savings. A preliminary list, Kammen says, includes research centers for energy efficiency, solid-state lighting, carbon sequestration, and green buildings, and a policy center to mesh California's climate regulations with those of other states and countries.

California needs technological advances

to meet its new greenhouse gas emission standards. The first standard, enacted last year, requires a 25% reduction in greenhouse gas emissions, to 1990 levels, by 2020. Two years ago, Governor Arnold Schwarzenegger signed an executive order targeting 80% reductions by 2050.

The new institute is not expected to fund new buildings, Kammen says, but rather will support and extend existing campus research efforts across the state. "Coordination is key here," says Ellen Auriti, executive director for research policy at UC's Office of the President. A public comment period closes next week, followed by hearings early next year. If all goes smoothly, the institute could have money to spend by next summer.

—ROBERT F. SERVICE

AIDS RESEARCH

Reconstructing the Origins of the AIDS Epidemic From Archived HIV Isolates

Five HIV isolates that had been forgotten in freezers for 2 decades are revealing new details about how and when the virus spread from Africa to Haiti and then exploded on the world scene. Evolutionary biologist Michael Worobey of the University of Arizona in Tucson led the new study, which analyzed HIV saved from five Haitian AIDS patients treated in Miami in 1982 and 1983. "It was the next best thing to being able to travel back in time," says Worobey, who obtained the samples through the U.S. Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia.

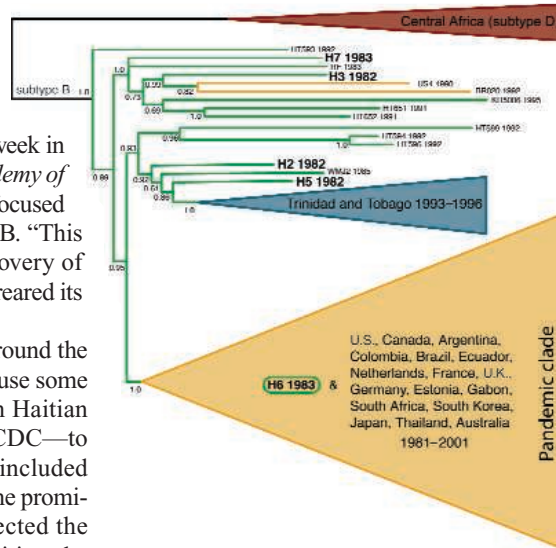
In a paper published online this week in the *Proceedings of the National Academy of Sciences*, Worobey and co-workers focused on what's known as HIV-1 subtype B. "This was the variant that led to the discovery of AIDS and so much of the story that reared its head after 1981," says Worobey.

Much controversy has swirled around the origins of the AIDS epidemic. Because some of the first AIDS cases surfaced in Haitian immigrants to the United States, CDC—to the consternation of many—once included Haitians as a special risk group. Some prominent Haitian researchers have rejected the idea that the virus spread from Haiti to the United States, contending that it likely moved in the other direction.

Molecular analyses of the archival isolates confirmed earlier reports that subtype B traveled from central Africa to Haiti about 1966, entering the United States 3 years later. The researchers' estimated probability that the virus instead traveled from the United States to Haiti—0.00003—is infinitesimal. "The methods are beautiful, and the analysis is elegant," says Bette Korber, an immunologist at Los Alamos National Laboratory in New Mexico, who published similar results in *Science* in 2000 (9 June, p. 1789).

Some are not persuaded. Jean "Bill" Pape, who heads the largest AIDS research program in Haiti, says Worobey and co-workers simply "restate prejudices advanced 2 decades ago." Pape notes that the authors offer no details about the sexual histories of the five Haitian immigrants, who he contends could have been infected by Americans. He also questions whether HIV arrived in 1966, pointing to retrospective studies in Haiti that did not find an AIDS case until 1978.

Other AIDS researchers counter that the Worobey paper offers the clearest picture yet of how the young epidemic matured. "It's a very nice piece of evolutionary sleuthing," says Beatrice Hahn, a virologist at the University of Alabama, Birmingham, and a co-author of the Korber study. One provocative finding, says Hahn, suggests that although several different isolates of subtype B came from Haiti to the United States, only one got a



Descent of HIV. A new analysis of stored blood samples from early AIDS patients shows Haiti (green) as a steppingstone between central Africa and the rest of the world.

foothold. It had not evolved ways to transmit more readily, says Worobey, and appears to have been "lucky" to have spread among high-risk populations—primarily, gay males in the United States. It then spread to Canada, South America, Europe, Asia, and even back to Africa (see figure).

Anne-Mieke Vandamme, a molecular epidemiologist at the Rega Institute for Medical Research in Leuven, Belgium, and co-author of a 2003 study that arrived at similar conclusions, says the new work underscores a fundamental feature of HIV epidemiology. Most of the early isolates found in Haitians quickly "died out," she notes. "You need an event that boosts the transmission, and the epidemic takes off." In this case, Vandamme says the promiscuity of gay men appears to have boosted the prevalence above a threshold that allowed the virus to thrive.

—JON COHEN

China Wants More Enviros ...

BEIJING—At last month's Communist Party Congress, China's leaders enshrined environmental protection in the country's constitution. Now China's State Environmental Protection Administration (SEPA) has inked a deal to train grassroots conservationists. SEPA's China Environmental Culture Promotion Association and Rare, a conservation group in Arlington, Virginia, will train budding Chinese conservationists in techniques—such as festivals and puppetry—that can stir public interest and pride in biodiversity in order to "translate knowledge into personal, meaningful change," says Brett Jenks, president of Rare. Southwest Forestry University in Kunming City in China will help launch projects at 10 sites next year, most likely in some of China's roughly 2000 nature reserves.

—RICHARD STONE

... And Heads to the Moon

The first spacecraft launched beyond Earth orbit by a developing nation is on its way to the moon. Chang'e 1, named for the Chinese goddess who flew to the moon, will arrive in lunar orbit 5 November. The 24 October launch drew large crowds near the Sichuan launch center, was broadcast live on national television, and prompted senior Chinese officials to declare plans to share culled science data. The 2300-kg satellite will circle the moon for a year and send back three-dimensional images of the lunar surface and an analysis of moon dust. India and the United States plan to launch moon orbiters next year, and Japan announced this week that it will launch a robotic rover in the next decade.

—ANDREW LAWLER

Oceans Are Nickel-and-Dimed

HONOLULU—A dozen marine scientists gathered here last week at the behest of the International Seabed Authority to design safeguards against the anticipated damage from the industrial harvesting of potato-sized nodules rich in nickel and copper sitting on a part of the Western Pacific sea floor with great biodiversity. "Practically every individual [organism] is a new species," said Alex D. Rogers of the Zoological Society of London. The scientists inserted a patchwork of nine 400-km-by-400-km protected areas, in between mining claims in an area nearly the size of Australia. Harvesting is expected to start within a decade. If adopted, as expected, the restrictions would be the first such sanctuaries in international waters.

—CHRISTOPHER PALA

HISTORY OF SCIENCE

Scientists Fete China's Supreme Polymath

SHANGHAI—In the early 17th century, this humanist and experimentalist helped avert starvation in China by disseminating hardier crops and devised dams and canals for irrigation and flood control. He launched a decade-long effort to improve the accuracy of the Chinese calendar by incorporating a more precise knowledge of celestial geometry. His monumental contribution was to team up with a Jesuit scholar to translate part of Euclid's *Elements*, introducing late Ming Dynasty intellectuals to new mathematical concepts—and Western thought. For his achievements, he has been compared to Leonardo da Vinci and Francis Bacon.

Who was China's Renaissance man? Go to the head of the class if you guessed Xu Guangqi.

Last month, scientists from a variety of disciplines met here at the Partner Institute for Computational Biology (PICB) to commemorate the 400th anniversary of the publication of the first six volumes of *Elements* in Chinese and to explore Xu's remarkable legacy. "He started China's enlightenment," says cell biologist Pei Gang, president of Tongji University in Shanghai. "Xu promoted the idea of learning from the West." Over the past century, Chinese leaders have taken Xu's advice to heart, including a reference by President Hu Jintao at last month's Communist Party's 17th National Congress to the importance of taking a "scientific view of development."

Xu (pronounced like "sue") was born in Shanghai in 1562 and was groomed to be a civil servant. A watershed moment came in 1600, when Xu met Matteo Ricci, an Italian Jesuit and one of the first Westerners allowed to live in China. No intellectual slouch himself, Ricci had been a student of Europe's leading mathematician of the time, Christopher Clavius. "Xu was wholeheartedly attracted to him," says Hung-lam Chu, a historian at the Chinese University of Hong Kong. The kindred spirits came to realize that planar geometry and other higher mathematical concepts then unknown in China were essential to progress. "*Elements* is basically a book about Western logic," says Yu Sanle of Beijing Administration Institute.

In translating *Elements*, Xu and Ricci also coined a host of terms, including *jihe* as the character for "geometry." Ricci deserves equal billing, Yu argues: "His was the greatest contribution of any foreigner to Chinese culture and civilization." Knowledge flowed both ways: Ricci also translated several



Leading light. Among many contributions to China, Xu Guangqi helped disseminate Western knowledge of geometry; title page of Euclid's complete *Elements*.



Confucian classics into Latin.

After earning a *jinshi* degree, the equivalent of a Ph.D., in the palace examination of 1604, Xu was admitted to Beijing's prestigious Hanlin Academy. He ascended smoothly through the government ranks until late in his career he came to be known, simply, as "The Minister." Throughout his life, one constant was his dedication to improving agriculture. His experiments in Shanghai with yams, then a new import from South America, led to the widespread adoption of the high-energy crop. "This was decades before the West began taking a scientific approach to agriculture," says PICB director Andreas Dress. Xu also trained imperial soldiers to use a newfangled device from Europe, the cannon. "He was a fascinating polymath who spread his interests far and wide for a specific purpose: statecraft," says Dagmar Schäfer of the Max Planck Institute for the History of Science in Berlin, Germany.

Xu was also a key figure in China's calendar reform. China's calendar was based on observed motions of the sun and moon, whereas the West's was based on average motions. "The Jesuits had better data than Late Ming astronomers and a clear geometry

of the heavens," says Peter Richter, a theoretical physicist at the University of Bremen in Germany. With such knowledge, Ricci predicted that an eclipse would occur on 15 December 1610—right on the money. "That impressed Xu Guangqi," Richter says.

After Xu and Ricci's successors correctly predicted an eclipse in 1629, the emperor appointed Xu as leader of the calendar reform, which he embarked on with the assistance of Jesuits. The reform was completed after Xu's death in 1633. The reams of data used to justify the revision amounted to the first scientific collaboration between scientists in Europe and the Far East.

Xu's tomb is a 20-minute drive from PICB, jointly run by the Chinese Academy of Sciences (CAS) and the Max Planck Society. "That's how we learned about this guy," says Dress, a computational biologist with an abiding interest in science history. "Suddenly it dawned on me" that 2007 is the anniversary of Chinese *Elements*. Accordingly, Dress broadened the workshop into an International Xu Guangqi Conference, organized by CAS, the Shanghai Institutes for Biological Sciences, and Shanghai Xuhui District Government.

One aspect of Xu's remarkable life that the Chinese government rarely talks about is his conversion to Roman Catholicism and baptism in 1603 as Paul Xu Guangqi. Although some scholars argue that Xu converted out of gratitude or in recompense for Ricci's help, Chu and others are convinced that he was both a devout Christian and a faithful Confucian, noting that Xu's writings consistently adhere to Confucian philosophy. Nor were the Jesuits as zealous in their missionary work as groups that came later: After all, notes Richter, Ricci helped translate *Elements*—not the Bible—into Chinese.

Xu's legacy was imperiled by the collapse of the Ming Dynasty in 1644. But his seminal contributions are cherished by modern China. "He forged a dialogue between the West and China," says Li Tiangang of the Shanghai Academy of Social Sciences, a dialogue that grows richer, and more relevant, by the day.

—RICHARD STONE



Who Owns Glycobiology?

A group of biologists who claim to be cleaning up their field are battling a Texas company that makes glyconutrients; the company, threatening legal action, says these scientists are spreading falsehoods

MOST BASIC BIOLOGISTS PROBABLY don't give much thought to ginseng, herbal teas, and the other dietary supplements that fill the shelves of health-food stores. But one group has been caught up in an escalating controversy over supplements in recent weeks, pitting them against a nutraceuticals company that has been charged with illegal sales practices in Texas. Several prominent U.S. glycobiologists allege that the company, Mannatech, threatens to "taint" their field by linking glycobiology discoveries with claims for sugar pills of unproven medical benefit. The company and its consultants have fired back, urging the withdrawal of a critical commentary and warning of legal action.

The dispute began when a leader in the field and the editor of the journal *Glycobiology* co-authored a scathing critique of Mannatech, which was posted as an online preprint in September. Other scientists joined in, urging speakers to reconsider their participation in a scientific meeting last week in Dublin that was partly sponsored by Mannatech. In

another twist, after Mannatech and others complained about the *Glycobiology* editorial, the publisher, Oxford University Press, removed it 3 weeks ago, promising to repost it later with responses.

Mannatech Inc., a publicly traded company in Coppell, Texas, with more than \$400 million in sales last year, was under fire from legal authorities before this controversy broke. In July, the attorney general of Texas charged the company with violating Texas laws and has sued to stop its sales associates from making claims that its products cure diseases. The company has denied the charges but has said that it is revising its literature.

Independently, several U.S. glycobiologists say they want to set the record straight and prevent damage to their field. "It's none of our business as glycobiologists if nutraceutical companies want to sell bark extract. When they begin to tie it to our discipline, that is the problem," says Ronald Schnaar of Johns Hopkins University in Baltimore, Maryland, editor-in-chief of

Glycobiology and lead author of the withdrawn editorial. Adds Ajit Varki, a glycobiologist at the University of California, San Diego, this "striking example" illustrates "the responsibility of scientists to speak out."

But others, including several European glycobiologists, suggest that these criticisms are overreaching. The chief organizer of the Dublin conference, John Axford—a British rheumatologist and until recently a Mannatech board member—says he and other organizers were "flabbergasted" by the letter suggesting speakers not attend the conference. He says Mannatech's support is no different from the drug company money that funds countless scientific meetings. His camp argues that the Mannatech critics are unfairly dismissing the entire complementary medicine industry, which should be encouraged to take part in scientific meetings, not banned from them.

Sugar therapy

The science at the heart of this dispute concerns not dietary sugars but sugar

Sweet treatment. Mannatech's flagship product, Ambrotose Complex, is a blend of plant extracts containing polysaccharides.

molecules linked to most proteins and many fats in the body. These glycoproteins and glycolipids dot the surface of cells and serve as identification markers recognized by other cells. The field took off in the past 2 decades, and carbohydrates are now widely studied for their roles in normal biology and disease (*Science*, 23 March 2001, p. 2337).

Mannatech, started by entrepreneur Samuel Caster, who had previously run businesses selling pest-control devices and home insulation, launched its first product in 1994. That year, Congress passed a law allowing companies to sell dietary supplements without Food and Drug Administration approval, as long as they don't claim any therapeutic benefit. The company's Web site states that its glyconutrient products "apply concepts from the rapidly emerging field of glycobiology." A press release goes on to say that eight sugars "are needed at the cellular level for optimum immune function" but that "six of these dietary glyconutrients ... are often lacking in our modern diets." The company's main product, Ambrotose Complex—a mixture of larch bark arabinogalactan, aloe vera gel, gum ghatti, and gum tragacanth—is "a blend of specific plant saccharides that support the immune system" and "support optimal cell-to-cell communication," a product data sheet says.

Mannatech relies on more than 500,000 independent salespeople, or associates, to sell Ambrotose. According to a lawsuit filed in July 2007 by the Texas Attorney General, the company's sales associates promote its products with testimonials from people who claim glyconutrients have cured diseases such as Down syndrome and cancer. The company also "claims scientific validation from the field of glycoscience" for its products, the Attorney General concluded. Mannatech said in response that it "intended[ed] to cooperate to reach a resolution" with the state; meanwhile, the Attorney General's office says that it is pursuing the charges in court.

In their commentary, Schnaar and co-author Hudson Freeze of the Burnham Institute for Medical Research in San Diego,

California, examine Mannatech's educational Web site, GlycoScience.org, shuttered since August for revisions related to the Attorney General suit, Mannatech says. The archives include a self-described "peer-reviewed scientific journal" called *Glycoscience & Nutrition*, edited by Mannatech scientists and consultants. Its articles describe glycobiology findings "mixed with speculation about the potential of glyconutrients to positively impact" various disorders, write Schnaar and Freeze. Case reports on diseases such as breast cancer and cerebral palsy treated with glyconutrients also appeared in the *Proceedings of the Fisher Institute for Medical Research*, which is published by an independent organization described by the Texas Attorney General as "little more than a sham charity with the sole purpose of providing 'scientific' support to the illegal health claims made about Mannatech's products."

As for peer-reviewed research on the health benefits of Ambrotose in PubMed-listed journals, there's not much, according to Schnaar and Freeze. Their review of the literature in the past 25 years found that larch bark arabinogalactan isn't digested, and its main effect is flatulence. About 30 in vitro and animal studies on ingesting aloe gel glucomannan, largely funded by Mannatech or its supplier company, suggest some immune system effects, but the only clinical trial in a PubMed-listed journal found no benefit. The notion that any saccharide mixture will improve health is flawed, says Varki, because most people's cells can make all the sugars they need from glucose in the diet.

Despite the scant support for glyconutrients in mainstream scientific publications,

many glycobiologists have found their names on Web sites promoting Mannatech products. References to Freeze's papers appear on Mannatech sites because 12 years ago, he showed that ill children with a rare sugar-processing disease recovered if they were fed mannose. Ever since, Freeze says he has received calls and e-mails from families of patients with diseases such as cancer and autism. They want to know if spending \$34, \$100, or more per month on Ambrotose will help. James Paulson of the Scripps Research Institute in San Diego, California, began getting calls after he became head of a glycobiology consortium funded by the National Institutes of Health (NIH) in 2001. "It just tore me up knowing my name was being used in that way," he says. Three Nobel Prize winners at Rockefeller University—including Günter Blobel—filed a complaint against Mannatech last year to make the company and associates stop using their names on Web sites. The company "has accepted some responsibility" to patrol such Web sites, says Rockefeller general counsel Harriet Rabb.

Particularly disturbing, say Schnaar and Freeze, are references to the benefits of glyconutrients in a medical textbook called *Harper's Illustrated Biochemistry* from McGraw-Hill, whose lead editor is a retired University of Toronto biochemistry professor, Robert Murray. Murray, who is also a Mannatech consultant, added a paragraph to the 2006 edition noting that although the body can make other monosaccharides from glucose, "there is evidence that the other sugars may be beneficial in some circumstances when added to the diet. This has led to the development of glyconutrient supplements,"

which are "under study," the text notes. Schnaar and Freeze say this statement, which Mannatech cites in a fact sheet, could mislead young scientists and doctors.

Murray responds that "there is some scientific basis to the way [Ambrotose] affects wellness," noting that "thousands of people claim to have benefited" from it. He says he has done nothing wrong and has demanded through a lawyer that his name be removed from the editorial by Schnaar and Freeze.

The Jenner meeting

The most recent example of Mannatech's tactics, critics say, is its sponsorship of the Jenner Glycobiology and Medicine



Nutraceuticals heavyweight. Mannatech Inc., headquartered in Coppell, Texas, had sales last year of more than \$400 million.

Symposium, a biennial meeting organized by the Royal Society of Medicine (RSM), a U.K. educational organization. Axford, a member of the RSM board and head of a research center at the St. George's University of London, has organized the meeting since it began in 1991. Axford served on Mannatech's board from 2002 until he stepped down in September; he has received at least \$420,000 in compensation, \$370,000 in stock, and \$200,000 in research contracts from Mannatech since 2000, according to reports filed with the U.S. Securities and Exchange Commission.

After the 2004 Jenner meeting, Mannatech posted on its GlycoScience.org site a summary by Axford with an added phrase stating that the meeting's "major sponsor was Mannatech, Inc." Last year when Axford proposed holding the meeting in San Diego, some glycobiologists objected. According to Scripps's Paulson, he and other local organizers asked RSM to return \$30,000 in support from Mannatech, and, after that, the meeting was canceled for lack of funding. Axford says the meeting fell apart because of low registration and a failure to find local sponsors.

This year, the controversy blew up again when Schnaar and Freeze decided to send the speakers invited to the Dublin meeting a copy of their critical editorial along with a letter signed by themselves, Paulson, Varki, and Johns Hopkins's Gerald Hart. The letter alleges that Mannatech sites have "prominently displayed" the Jenner meeting and that "your participation may be misused to impress, mislead, or confuse desperate and vulnerable people."

It urges speakers to make sure the meeting is not sponsored by Mannatech and take "whatever actions you feel are appropriate."

The letter apparently had little impact. Several speakers told *Science* that RSM assured them that Mannatech was providing an unrestricted educational grant and that it would have no influence on the content of the meeting. RSM "set up the meeting in a responsible and sensible manner," says Anthony Corfield of Bristol University in the U.K. RSM education director Jo Parkinson wrote *Science* that Mannatech is just one of five company sponsors, and that RSM policy "in no way allows the sponsoring company to claim any association with the RSM." Axford told *Science* that he agreed that Mannatech

"It's none of our business as glycobiologists if nutraceutical companies want to sell bark extract. When they begin to tie it to our discipline, that is the problem."

—Ronald Schnaar,
Johns Hopkins University

natech calls the editorial "a deliberate attack ... based on false statements." The company, which provided *Science* with a copy of the letter and a response to *Glycobiology*, cites more than two dozen studies, including human studies, it has published or presented at meetings and says it has spent more than \$20 million on research.

Axford defends his involvement with the company, saying he's "helped them turn

Gray literature. Many clinical case reports and reviews of the health effects of glyconutrients have appeared in two journals that aren't listed in PubMed.

could post his summary of the 2004 Jenner meeting on its Web site for educational purposes. He says he had overlooked until now that Mannatech had added a phrase stating that it was the "major sponsor," which was incorrect, he says. Parkinson says there was no violation of RSM policy.

The editorial vanishes

Meanwhile, Oxford University Press has removed the Schnaar-Freeze commentary and plans to post it later with responses. Biosciences senior editor Claire Bird says Oxford Press had received complaints from Mannatech, Axford, and Murray that the

article was not properly peer-reviewed. The authors and the press have agreed the posting should be revised. Schnaar says he hopes for "a real forum with space for all interested parties." Bird says guest editors will oversee the submissions.

In its letter to Oxford Press, Man-

atech calls the editorial "a deliberate attack ... based on false statements." The company, which provided *Science* with a copy of the letter and a response to *Glycobiology*, cites more than two dozen studies, including human studies, it has published or presented at meetings and says it has spent more than \$20 million on research.

atech calls the editorial "a deliberate attack ... based on false statements." The company, which provided *Science* with a copy of the letter and a response to *Glycobiology*, cites more than two dozen studies, including human studies, it has published or presented at meetings and says it has spent more than \$20 million on research.

atech calls the editorial "a deliberate attack ... based on false statements." The company, which provided *Science* with a copy of the letter and a response to *Glycobiology*, cites more than two dozen studies, including human studies, it has published or presented at meetings and says it has spent more than \$20 million on research.

Even Mannatech's critics agree that research on the potential health benefits of consuming specific sugars has merit. Varki, for example, says he is interested in sialic acids as treatments but feels "inhibited" because of the stigma attached. If the field of glycobiology is to expand, many leaders agree, health claims must be based on solid science.

—JOCELYN KAISER



NUCLEAR PHYSICS

A Lab to Get the Measure of Matter

Germany's GSI laboratory aims to become Europe's premier nuclear physics center with a set of accelerators and detectors that will be the envy of the world

DARMSTADT, GERMANY—Although it cannot yet claim instant name recognition, a German national lab in this small town near Frankfurt has earned its 15 minutes of fame several times over. GSI, as it's called, has forged new chemical elements never before seen on Earth. It celebrates its own successes by presenting visitors to the lab with a packet of six candies in different colors. The first is labeled "Bohrium" and bears the date 24 February 1981, the day on which GSI's nuclear alchemists created this element with atomic number 107. The last represents element 112, discovered in 1996 but still unnamed as it awaits verification.

Although a rival lab has added five other superheavy elements to the periodic table since then, GSI hopes to soon grab the headlines again, in many areas of research. Over the next decade, the lab will more than double in size with the construction of two high-energy superconducting synchrotrons capable of accelerating any atom up to uranium, three storage rings to condition particle beams, and numerous new detectors serving five different fields of physics. Construction of the new additions, dubbed the Facility for Antiproton and Ion Research (FAIR), will be officially launched next week. FAIR will be an international affair, with more than a dozen other countries contributing to the

€1.2 billion price tag. "It's a big, new venture, a multidisciplinary facility with a very broad range of experiments," says Sydney Gales, director of France's heavy-ion laboratory GANIL in Caen. Those involved in designing the project are ecstatic. "Starting up such a facility is a once-in-a-lifetime chance," says FAIR physicist Thomas Beier.

With the wide array of high-energy particle beams available at FAIR, researchers will be able to study the particle soup that existed at the very beginning of the universe, the workings of atomic nuclei, the dense plasma in the cores of brown dwarfs and gas-giant stars, how stellar explosions create heavy elements, and whether antihydrogen differs from its normal form. Physicists here and elsewhere are already addressing these questions, but FAIR will provide particle beams with as much as 10,000 times the intensity and 30 times the energy of those now available at the lab, with the added bonus of antimatter beams that make new types of experiments possible. "This type of machine and energy is not available anywhere in the world today," says Gales. "I'm eager to see what they find."

Exotic studies

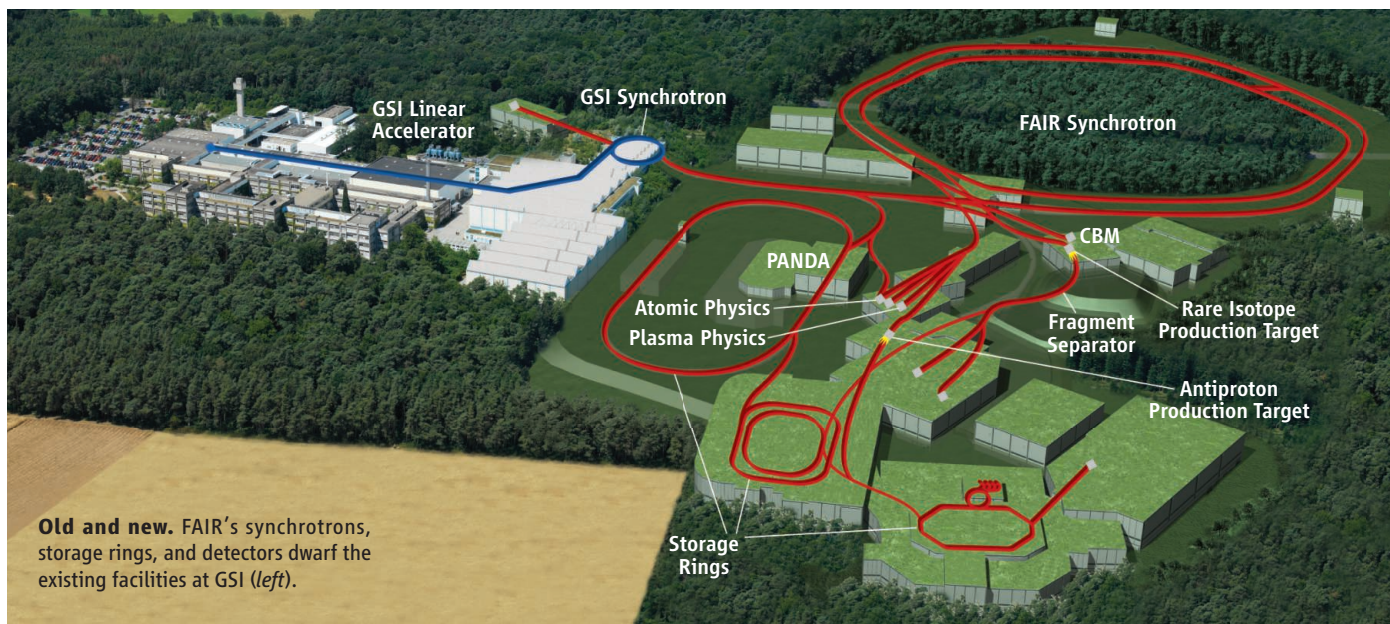
GSI was founded in 1969 as a shared resource for several universities in the

German state of Hessen. The lab started out with a single linear accelerator that could accelerate any element from hydrogen to uranium up to 20% of the speed of light. A synchrotron added later boosted speeds to 90% of light's.

In the late 1990s, GSI's researchers began discussions with the 1200 outside users on what the lab should do next. By 2001, the lab had drawn up a conceptual design report for a new facility that would expand on GSI's work with particle beams and add an ability to use antiprotons. Germany's federal government agreed to fund the plan if at least 25% of the costs were borne by foreign partners. So far, 13 partners apart from Germany have signed up, including European nations and Russia, India, and China. (The United States is working out its own plans for rare isotope research and is not involved.)

Construction of FAIR will begin in a year, and the facility should be complete by 2016. GSI will remain a German lab, and a new company set up under German law will run FAIR but be owned by the 14 international partners. This arrangement didn't go down well with some countries, particularly France, which preferred an international body like other pan-European research centers. Although this issue didn't prove a show-stopper, some countries have yet to confirm their financial contribution.

GSI and FAIR intend to work closely together, but all that GSI will physically provide to the new facility is a beam of particles already sped up by its linear accelerator and synchrotron. The beam will then be passed to FAIR's new synchrotrons to be boosted to higher speeds. After that, the resulting



particle beam will periodically smash into a target in a detector that will analyze the debris to provide insight into the state of matter at the moment of collision. At other times, the beam will hit a target, and the debris will be collected into a new beam containing all sorts of nuclei. These will be sifted on the fly to pick out particular unstable nuclei for study.

These nuclei can be passed to a storage ring, which allows researchers to control the speed of nuclei and create a precise, high-quality beam for subsequent experiments. "FAIR will have two unique abilities," says nuclear physicist Donald Geesaman of Argonne National Laboratory near Chicago, Illinois, "high energy and storage rings, which allow you to do certain types of experiments."

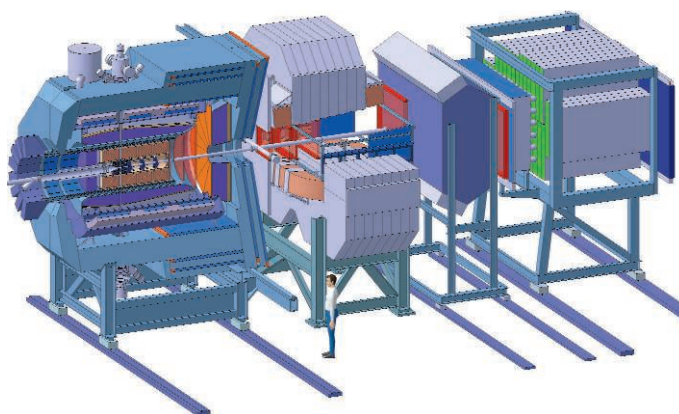
Planning for FAIR has involved about 2500 scientists and engineers worldwide. "We've never worked in such large numbers before," says FAIR physicist Ingo Augustin. Although it serves several fields, GSI's biggest program is in exotic nuclei. FAIR will have a similar emphasis: Some 700 scientists from more than 30 countries devised FAIR's exotic-nuclei research program. Exotic nuclei are highly unstable isotopes that are key to finding out the origin of many of the chemical elements on Earth.

Although a few elements were made by the extreme temperatures and pressures that existed soon after the big bang, many more—up to iron and nickel on the periodic table—are forged in the cores of stars. The heaviest elements have an even more violent birth in stellar cataclysms such as supernovae, gamma ray bursts, or x-ray bursts. In such explosions, lighter elements are bombarded with neutrons or with protons, swelling up to a large size but becoming increasingly unstable. Understanding how these unstable nuclei decay into something more long-lasting should tell physicists about the origins and abundance of the elements around us now. "If we know the nuclear physics, it constrains the astrophysics. Both are not well known now. This will change dramatically with FAIR," says GSI research director Karlheinz Langanke.

Other labs are chasing some of the same quarry, including CERN in Geneva, Switzerland, France's GANIL, Japan's new Radioactive-Isotope Beam Factory at the RIKEN lab in Wako, and the Facility for Rare Isotope Beams being developed in the United States. Although there will be overlaps with FAIR, the various facilities will

ultimately deploy different techniques, creating complementary research programs. "In each main region of the world, there are major investments in this area of research because people are excited about the science," says Geesaman. "None [of the facilities] can tackle all the problems coming from this large community."

FAIR is not all about nuclei. Atomic physicist Thomas Stöhlker is looking forward to more of the atomic dissection he does now



Back to basics. With a beam of antiprotons, FAIR's PANDA detector will probe the inner workings of protons, neutrons, and related particles.

at GSI but expects that FAIR will provide him with greater accuracy in nailing down atoms' properties. For example, FAIR should allow him to take a beam of uranium atoms, strip them of all their electrons, and then measure the mass of the naked nuclei with extreme precision, giving a better fix on electrons' binding energy.

FAIR will also provide the world's most intense antiproton beams, allowing researchers to better study the spectroscopy of antihydrogen and compare it with normal hydrogen. Any differences might help solve the long-standing mystery of why there is so little antimatter in the universe. FAIR will offer a "qualitative jump" in what's possible because of its dedicated storage rings and decelerator, says Stöhlker.

PlayStation, anyone?

Researchers hope to step out into entirely new terrain with antiproton beams, which FAIR will be able to produce and handle as at no other lab. They plan to study the strong nuclear force and the structure of hadrons—particles such as neutrons and protons that are made of quarks. Current theory does not explain the particular masses or spins that hadrons have. The three quarks that make up a proton, for example, account for only 10% of its mass and 40% of its spin. According to GSI hadron physicist Klaus

Peters, the main goal when FAIR is online will be to bombard hadrons with anti-protons to probe the binding of the quarks. The particles that mediate the strong force, known as gluons, "make most of the mass," he says.

Peters expects FAIR to become the most important lab in hadron physics: "There's no real competition." About 400 hadron researchers from around the world—"almost half of the community," Peters says—helped plan FAIR's program.

FAIR also has its sights on another prize: a quark-gluon plasma (QGP). This state of matter exists under such extreme conditions that quarks and gluons move almost freely in a soup of particles rather than being confined in hadrons. It is thought that the early universe existed as this form of plasma 20 to 30 microseconds after the big bang. Physicists at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory on Long Island, New York, believe that they have already created a QGP, but they have not yet

detected the telltale phase transition from normal matter. Some researchers now think that at the high temperatures used at RHIC, there may not be a phase transition, just a gradual shift from normal matter to QGP.

"What is still missing is proof of this new state of matter," says GSI's Peter Senger. "RHIC has some indications but no smoking gun." FAIR will approach the problem from a different direction. Instead of RHIC's high temperatures, FAIR will use high pressures, creating the sort of QGP found in neutron stars and one that has a detectable phase transition. But spotting the QGP will be no mean feat. In the fleeting instant of a particle collision, the putative plasma spews out thousands of particles, only a tiny fraction of which may carry the signature of a QGP. Senger and his colleagues have calculated that they must log 10 million reactions per second, beyond the abilities of today's computer processors. They've investigated prototype processors destined for PlayStation video game consoles not yet on the market. "These are the fastest processors, and if you have thousands working in parallel, it will be possible," Senger says.

Finding the phase transition to a QGP is "the major goal in heavy-ion physics," says Senger. And who knows, confirming a new state of matter just might give physicists in Darmstadt a new reason to hand out candies.

—DANIEL CLERY



IMMUNOLOGY

Testing the Line Between Too Much and Too Little

Keeping peanuts and other risky foods from toddlers in the hopes of preventing allergy has been common practice for years. But is avoidance actually safer?

In Israel, the joke goes, a child's first three words are Momma, Dadda, and Bamba. The last, a wildly popular kosher snack of peanut-flavored puffs, shows a cartoon baby with a tuft of red hair on its packaging. Many infants eat Bamba when they're as young as 6 months old.

Bamba has become an unlikely poster child in the world of food allergy, where physicians are desperate to ease the soaring burden of hyperactive immune responses. In the past 10 years, peanut allergies—among the most common and lethal—have roughly doubled in both the United Kingdom and the United States, where they affect at least 1% of young children.

A British study 4 years ago reported that peanut-allergic children have a lower quality of life than those with diabetes, suffering anxiety everywhere, from riding the bus to cruising the supermarket. “The kids don’t go to camp like they should, they don’t go to birthday parties,” says Brian Vickery, a fellow in allergy and clinical immunology at Yale University. They and their parents live with “fear and tremendous anxiety” that unexpected exposure will trigger shock and even death.

For years, physicians in many countries have urged parents to avoid feeding high-risk

children peanuts until they are 3 years old, on the rationale that one cannot become allergic to a food without being exposed to it and that the immune and digestive systems of older children will be better able to tolerate the peanut proteins that can elicit reactions. But in what seems a contradiction, the countries that endorse this view—including the United States, the United Kingdom, and Canada—have seen an explosion of peanut allergies. Meanwhile in Israel, with Bamba melting in the mouths of babies too young to walk, the rate of peanut allergy in toddlers is 0.04%.

Is this association accidental, or can it tell us something about how food allergies develop? It's a question asked increasingly by allergists. Early this year, a British team launched the first trial aiming for a direct answer. Their study is recruiting 480 high-risk babies and offering peanuts to half of them while withholding them from the rest. Will one group be more allergic? There's no safe bet, as studies of food, pet, and other allergies span the spectrum, with conflicting results on whether exposure is preventive.

Rethinking avoidance

An intricate set of variables can shape the onset of allergies: genetics, eczema, not

◀ **Eat or avoid?** Peanut allergies are on the rise, but steering clear of them early in life may or may not help keep them at bay.

being breast-fed, and the route of exposure, whether by the skin or the mouth. Some allergists wonder if food preparation plays a role—in particular, if the boiled peanuts consumed in Southeast Asia and Africa are less allergenic than the dry-roasted form eaten in the West. The surge in food allergies also tracks an increase in asthma and autoimmune disease in Europe and North America, and all are thought to be related to the plunge in infectious disease in early childhood. The “hygiene hypothesis,” as it's known, postulates that exposure to pathogens trains the immune system to regulate itself.

Given this complexity, there's little hope that avoidance will have universal effects. And indeed, the evidence so far is mixed. For some children, living with a cat may make cat allergies less likely, whereas living with more dust mites either helps or hurts as far as allergy prevention is concerned, depending on the study.

When it comes to peanuts, many physicians are torn between how little they know and what they viscerally fear: Data favoring early avoidance are skimpy, but doctors are uncomfortable promoting peanut consumption. “Peanut allergy is so lethal; you don't get a second chance,” says Patrick Holt of the University of Western Australia in Perth, who examines mechanisms underlying allergy. Still, he says, “the story that we thought we were seeing 10 years ago,” with allergy risk tracking a dose curve, “doesn't hold up.”

The thought that higher doses may not be worse has come gradually, as scientists dig deeper into how environmental cues prompt an immune response. That the immune system attacks something like *Escherichia coli* bacteria, a sickness-inducing pathogen, “makes sense,” says Stephanie Eisenbarth of Yale University School of Medicine, who performed allergy studies in renowned immunologist Kim Bottomly's lab for her Ph.D. thesis. “But why would you want to fight ragweed?” The common plant isn't a health threat. This is the question allergy experts have asked themselves for decades.

In someone vulnerable—because of their genes or for other reasons—initial exposures to normally innocuous proteins stir white blood cells known as T helper 2 (T_H2) cells into a tizzy. They alert B cells, which churn out antibodies to the protein. Those antibodies, part of a class called immunoglobulin E (IgE), bind to mast cells and set off a full-blown reaction the next time the offending

protein, called an allergen, appears. The mast cells empty inflammatory chemicals such as histamines into the bloodstream, causing the classic symptoms of allergy.

In the mid-1990s, Bottomly, now president of Wellesley College in Massachusetts, began reporting a more nuanced picture in cells and in mice: The immune response to a foreign substance shifted as the dose of that substance increased. At low doses, her lab saw a classic T_H2 response. High doses set off the flip side of the immune system, activation of T_H1 , which served as a brake on T_H2 . In 2002, Eisenbarth and Bottomly revealed in mice that the same dose pattern held with an adjuvant, a substance in the environment that primes the immune system to respond to an allergen. “We still don’t know” what’s behind this, says Eisenbarth.

Testing whether something similar occurs in people, some allergy-prevention studies uncovered a bell-shaped curve, particularly with exposure to pets: no reaction at the zero-dose end, a strong reaction with low or moderate doses, and little or no reaction again at the high-dose end.

Some foods may fall under this rubric, too. In 2004, pediatric asthma and allergy specialist John Warner of Imperial College London described a cohort of babies at high risk for egg allergy and examined the effect of maternal diet during pregnancy. Infants whose cord blood had the lowest or highest levels of antibodies to egg proteins were least likely to experience allergic hypersensitivity, such as eczema, by the time they were 6 months old. Those with midrange antibody levels were the most hypersensitive. The evidence is suggestive, although it remains unclear whether a pregnant woman’s diet can influence allergy in her unborn child.

Still, studies such as this one have researchers questioning “whether we can really prevent exposure” completely, says Dennis Ownby, a pediatric allergist at the Medical College of Georgia in Augusta. At least half the toddlers he sees who have suddenly reacted to a food weren’t known to have eaten it before. But because an allergy can’t develop without exposure first, they must have encountered proteins somewhere. Foods not known to contain peanuts may be one place: In 2001, the U.S. Food and Drug Administration reported that 25% of foods surveyed that did not list peanuts on their labels tested positive for them.

“The problem is the small amounts may be even more allergenic than the large amounts,” says allergist Hasan Arshad of the University of Southampton, U.K. No one

knows whether this is true in humans. But in cells and animals, the allergy antibody, IgE, can be stimulated with small doses of certain allergens given intermittently, whereas a protective antibody, IgG, comes forward with large doses given regularly, says Allan Becker, a pediatric allergist at the University of Manitoba in Winnipeg, Canada. Arshad’s own work, however, underscores the tension between avoidance and exposure.

The Isle of Wight experiment

Seventeen years ago, Arshad began following 120 infants on the Isle of Wight, a semi-rural island several kilometers off the coast



Omnipresent. Scientists are testing whether exposure to peanuts, dust mites (above), and grass can help prevent allergies in high-risk toddlers.

of England. The children, all at high risk of allergy and asthma, were put in a limited-exposure group or one with normal exposures. In the low-exposure group, breastfeeding mothers avoided cow’s milk, eggs, nuts, and wheat. Exposure to house dust mites was reduced with pesticides and mattress covers.

In their first 8 years, these children were half as likely to suffer allergies as the control group. Limited exposure worked on the Isle of Wight, Arshad theorizes, because several allergies were targeted simultaneously and because the children steered clear of small

exposures that may be deleterious, something other avoidance studies may not have accomplished.

Some physicians are looking beyond the level of exposure and considering how food proteins get inside the body. In 2003, pediatric allergist Gideon Lack, now at King’s College London, reported in *The New England Journal of Medicine* that by the time they were 6 months old, 21 of 23 preschoolers with peanut allergy had had creams containing peanut oil, which are typically used for diaper rash, applied to their skin. In a control group of 140 children, the number was 59%. Lack believes the route of exposure is key—specifically, that skin exposure in the absence of ingestion, rather than ingestion of small doses, is behind growing rates of peanut allergy. “It would be highly implausible to me that we’ve evolved as a species so that if we have tiny amounts of food through the gut,” an allergy develops.

One effort to settle the question comes in a study Lack launched early this year. Funded by the U.S.-based Immune Tolerance Network (ITN), it is randomizing babies younger than 11 months old to receive peanuts in different forms, such as peanut butter mixed with banana, or no peanuts until they’re 3 years old, and track who develops allergies. About 200 children have enrolled so far; the study will last 7 years.

Holt, in Australia, is running a similar ITN study of grass, cat, and dust-mite allergy, recruiting 200 high-risk babies and giving half of them daily doses of allergen drops under the tongue. “For people on the outside, this sounds like very adventurous stuff,” says Holt. But “any early fears of ‘Oh my God, you’re going to create havoc’” by exposing high-risk babies to allergens haven’t yet come to be. Results are still a few years off.

In both the United States and the United Kingdom, physicians and health officials are reconsidering guidelines that encourage parents to avoid giving high-risk children potentially allergenic foods; they may shift to a more neutral stance. But “I’m still nervous about peanut,” admits one leader in the field, Hugh Sampson of Mount Sinai School of Medicine in New York City. “I don’t have proof, I just have this sort of sense that there’s something different about it.” Asked by worried parents what he would do were the child his own, “I’ll say, ‘I would probably avoid peanut.’” At the same time, Sampson admits to never quite knowing what to advise. “Whatever we’re doing is not working,” he says, “because things have only gotten worse.”

—JENNIFER COUZIN



ECOLOGY

Do Wandering Albatrosses Care About Math?

Repudiating a decade-old study of sea birds, a new report questions a popular model of how animals—as well as fishing boats and people—search for food

As its name suggests, the wandering albatross (*Diomedea exulans*) is a fabulous flyer, flapping its way around the world with the help of its 3-meter plus wingspan, the longest of any living bird. A 1996 report seemed to offer clear proof of the bird's endurance: It found that the birds sometimes soared for as many as 4 days before touching down on water, presumably to catch fish or other food. One of the first studies in which recording devices tracked animal movements, the work also brought a little-known mathematical tool to bear on the study of animal foraging. It showed ecologists that a model of random motion called a Lévy flight described the way albatrosses searched for food.

Inspired by the work of French mathematician Paul Lévy, Lévy flights are characterized by many short hops, with much longer jumps on rare occasions. Physicists have long used the mathematics behind Lévy flights to predict how particles move in liquids and how matter spreads in the universe, for example. And after analyzing the recorded albatross data, a team led by physicist H. Eugene Stanley of Boston University (BU) concluded in a 1996 *Nature*

article that the tagged birds also followed Lévy flights.

This strategy could work well when food supplies are concentrated in a few places—say, in schools of fish or fields of flowers—with long stretches of empty sea or bare ground in between. Indeed, in a 1999 *Nature* article, Stanley's group outlined the theoretical benefits of foraging with the strategy and asserted that deer and bumblebees followed a Lévy flight search pattern.

Last week, however, a new study revealed flaws in both the 1996 and 1999 reports. In the initial albatross paper, Stanley's team misinterpreted key data on the birds, says Andrew Edwards of the Pacific Biological Station in Nanaimo, Canada. In the reanalysis by Edwards and colleagues, reported in the 25 October issue of *Nature*, the longest flight drops from 99 hours to 20 hours, for example. Moreover, using an arsenal of statistical tests on the 1996 and 1999 papers, the researchers show that the bird, bumblebee, and deer data support other search strategies equally well. Edwards goes so far as to say that none of the subsequent studies reporting animal Lévy flights that he has seen are "100% convincing."

Stanley's team apparently accepts this dramatic about-face. All the surviving authors from the 1996 and 1999 papers are co-authors on the new report. And Gandhimohan Viswanathan of the Federal University of Alagoas in Maceió, Brazil, who was the first author of the two earlier papers, agrees with Edwards that "the jury is out" on whether Lévy flights apply to any foraging animals. "One message that this new paper sends is clear: We must be more careful with data analysis," he says.

Wet and dry birds

Physicists first discussed the notion that animals perform Lévy flight searches in the 1980s, but obtaining data on large-scale movements of animals in a natural environment wasn't easy at the time. When the 1996 albatross report came out, the idea of Lévy flights in animals really took off among ecologists. That publication has been cited more than 100 times. Researchers have since described Lévy flights in jackals, spider monkeys, seals, microscopic zooplankton, and even by fishing vessels and hunter-gatherer tribes.

All this excitement piqued Edwards's curiosity in 2005. An ecological modeler, he had just landed a position with the British Antarctic Survey (BAS) in Cambridge, U.K., the source of the albatross data analyzed by Stanley, Viswanathan, and their colleagues. Edwards decided to take a closer look at the original study to understand the methods used.

In 1992, BAS researchers had gone to Bird Island in South Georgia and attached a newly designed detector that registered contact with saltwater to a leg of each of five birds, retrieving the detectors about 2 weeks later at the birds' nesting sites. After a BAS group member met someone on the BU team interested in modeling animal movements, BAS agreed to provide the detectors' data.

For their analysis, the BU team interpreted "wet" signals as moments when the birds stopped flying over open sea and grabbed a snack from the water. The time in between was considered to be flight time, with longer "dry" intervals signifying longer distances in the air. Viswanathan, then a BU graduate student, Stanley, and their co-authors found a "hop, hop ... long jump, hop ..." pattern among the albatross journeys. When they graphed these data a particular way, the probability for a jump of a specific distance seemed to follow a so-called power-law distribution, in which most jumps were very short, and the longer a jump was, the rarer its occurrence. This type

of probability distribution fit the definition of a Lévy flight pattern.

Yet Edwards noticed that the first and last flights for each bird were suspiciously long. When he removed those flights from the original analysis, any evidence for Lévy flights vanished. The clincher came when, during a coffee break with a BAS albatross specialist, Edwards learned about data that the BU team didn't know about. For some of the recorded trips, some of the birds also wore rudimentary location trackers for a separate study. When Edwards and his colleagues at the survey pulled those data from the archives, they discovered that for much of the initial long "flights," the birds had never left their nests. Similarly, the last "flights" tended to be much longer because the birds had returned to their nests and were high and dry for hours. Flights recorded as 69, 67, 44, 26, and 23 hours in the original paper were actually all 4 hours or shorter, sometimes less than an hour.

Edwards and his colleagues also joined with the original BU authors to repeat the sea-bird study, drawing upon BAS data taken in 2004 from 20 Bird Island albatrosses equipped with more sophisticated tracking equipment. The new study confirmed that Stanley's team had been misled. "The birds aren't performing Lévy flights when foraging," Edwards says.

Edwards then took another look at the 1999 work on bumblebees and deer. He found other data of questionable relevance to Lévy flights: The deer "foraging" times analyzed by the BU team, for example, turned out to be the amount of time spent eating at a location rather than the time spent traveling between feeding sites.

Moreover, in both the 1996 and 1999 studies, the BU team had used a simple graphical approach to demonstrate the power-law distribution that signified Lévy flights, one that has since been applied by others in their analyses of animals' movements. But, says Edwards, the strategy can often produce a spurious conclusion of Lévy flights. Indeed, when Edwards statistically compared whether a power-law distribution fit the bumblebee and deer foraging data better than non-Lévy flight distributions, it didn't.

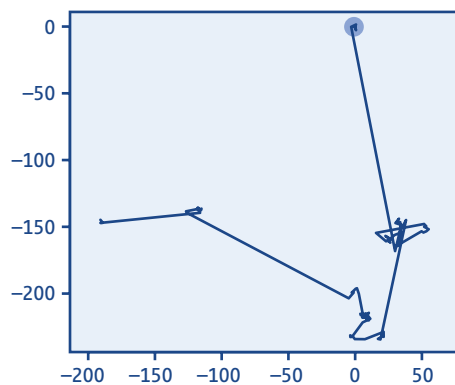
This type of sophisticated statistical model comparison, note Edwards and

Viswanathan, wasn't used in the 1990s. Still, the analysis removed from the two early papers the last bits of solid evidence for Lévy flights.

Those papers are not being retracted, however. Viswanathan says that he and the other original authors seriously considered retractions but decided that the new analysis serves as the needed public correction. Rory Howlett, an editor at *Nature* who handled the new paper, adds that its peer-reviewers didn't request retractions, either. "This is an unusual case in that new analytical methods and also new data became available that led to a re-evaluation of the original claims," he says.

From jackals to bees

Edwards's reappraisal of the two *Nature* papers isn't the only recent attack on biological Lévy flights. Behavioral ecologist Simon Benhamou of France's national research agency, CNRS, in Montpellier has also called into question the benefits of this foraging pattern. In the August issue of *Ecology*, he described computer simulations indicating that Lévy flight-styled movement is no more efficient at searching an area than is a variant of classical



Flight paths. Scientists have suggested that foraging jackals and bees (and even fishing boats above) follow a Lévy flight pattern (computer simulation, above), which is random motion marked by short spurts and occasional longer jumps.

Brownian random motion. Moreover, it takes a very close analysis of the data to distinguish the two types of movements. "Both my paper and Edwards's paper show that [the Lévy flight] concept is not likely to be useful to analyze foraging movements in most situations," says Benhamou.

To statistician Stephen Buckland of the Centre for Research into Ecological and

Environmental Modelling at the University of St. Andrews in Fife, U.K., the application of the Lévy flight concept to ecology was always "a bit far-fetched." "It's mathematicians taking a simplistic tool and pretending it is relevant to the real world," he says. The idea that foraging animals conduct Lévy flight searches, or any similar random search strategy, ignores that animals use their intelligence and experience to guide them, Buckland adds.

But not all ecologists share that dismissive view. They maintain that some Lévy flight studies may be valid. "My view is that the albatross work may well be flawed, but it has played a big role in stimulating some really good subsequent work on the whole issue of search strategies and animal behaviors," says Christopher Rhodes of Imperial College London, who has reported that foraging jackals perform Lévy flight searches.

Andy Reynolds of Rothamsted Research in Harpenden, U.K., is equally adamant that Lévy flights are a useful tool for ecologists. "Theory shows that Lévy flights are a good way to search. My feeling is that animals evolved to do the best possible searches," he says. Reynolds has found Lévy flight activity in starved flies searching for food in a container and in honeybees—tracked by radar—looking for food or hives in fields. He's also applied several statistical techniques to analyze his data. "I have used many methods to show Lévy flights rather than one method," says Reynolds.

Reynolds has been corresponding with both Benhamou and Edwards about his evidence for Lévy flights, but both remain unconvinced. Benhamou doesn't insist that ecologists abandon Lévy flights just yet. But he, like Edwards, argues that researchers must realize that there could be other, equally plausible explanations for how animals search: "Showing that data are well accounted for by a Lévy process is not enough."

—JOHN TRAVIS

Tiny but important
fishes

750



Planning for
future disasters

753



Regrowing body parts

754



LETTERS | BOOKS | POLICY FORUM | EDUCATION FORUM | PERSPECTIVES

LETTERS

edited by Jennifer Sills

Shopping for Explanations

WE WERE SHOCKED BY THE RANDOM SAMPLE "BORN TO SHOP?" (7 September, p. 1301). Such a study hardly deserves the notice of a premier science journal.

The Random Sample summarizes a "study" of sex-based differences in the ability of modern city dwellers to remember the locations of particular foods at a farmers' market. Given the social bias in American culture toward meal preparation by women and meal consumption by men, such a difference is not surprising, but neither is it indicative of "hard-wired" biological differences in brain function between women and men. Still less can such differences be attributed to an evolutionary past "when men were the hunters and women the gatherers."

While this gendered division of labor prevails (but is not universal) among ethnographically known foraging groups, such a pattern cannot be assum-



Response

BRUMFIEL *ET AL.* CRITICIZE C. HOLDEN'S summary (Random Samples, 7 September, p. 1301) of our research (1); we welcome the opportunity to respond.

Many studies document men's superior spatial performance (2). We found that women excel on a spatial task mimicking the cognitive demands of plant-food gathering, even when we used a measure that normally gives men an advantage. Brumfiel *et al.* suggest that ancestral sex differences in hunting may be small; however, this is irrelevant to our theory of gathering-related spatial adaptations. What is relevant is whether, statistically, ancestral women gathered more than men. If so, they could be the target of stronger selection for cognitive mechanisms supporting gathering. This sex difference in gathering is universal among described hunter-gatherers (3), and chimpanzee data suggest that it extends back to our pre-hominin ancestors (4).

Citing cultural biases in shopping and cooking, Brumfiel *et al.* present a social-learning explanation for our results. Their theory is contradicted by other studies and our

ed for the period of human evolution. Archaeologists argue that with the invention of spear-throwers, bows and arrows, and poison darts, Paleolithic hunting probably involved herd surrounds and game drives, such as those practiced by Native Americans in the western United States at the time of European contact. These surrounds and drives involved all camp members: men, women, and children.

An important trait that distinguishes modern humans from other species is their "hard-wired" ability to learn a wide variety of socially transmitted patterns of thinking and acting. Surely *Science* was premature in drawing public attention to a study that purports to say something about universal differences between women and men based on observations made in a single culture. Studies such as this reinforce the American inclination to explain our own culturally based gendered division of labor (e.g., women shop and cook, men are mathematicians) in terms of biology rather than patterns of socialization and discrimination.

ELIZABETH M. BRUMFIEL, MICAELA DI LEONARDO,
KATHERINE E. HOFFMAN, CHRISTOPHER W. KUZAWA,
THOM MCDADE, HELEN B. SCHWARTZMAN, REBECCA SELIGMAN

Department of Anthropology, Northwestern University, Evanston, IL 60208, USA.

data. First, studies show either no sex difference or a male spatial advantage in nonfood shopping environments (5). Second, counter

to the social-learning hypothesis, individual differences in shopping experience, taste preferences, and consumption frequency did not predict spatial performance in our study; women outperformed men controlling for these experience factors. Moreover, both sexes showed better performance on high-calorie food items. This is the signature of an evolved mechanism for efficient gathering, not one socially learned in contemporary environments.

Finally, only those who insist upon egalitarianism depend on claims of biological identity. The sexes differ. Men never gestate offspring. On average they are larger, less articulate, shorter lived, and better at mental rotation tasks (2, 6). Denying these and other differences will not make them disappear. But the science that explores these differences provides tools to combat

discrimination. For decades, researchers uninformed about our evolutionary history unknowingly constructed spatial tasks that favor men's skills. It is only when we take seriously men's and women's evolutionary heritages that we can break through this inadvertent sexism and expose women's unique abilities.

MAX M. KRASNOW,¹ DANIELLE TRUXAW,¹
JOSHUA NEW,² STEVEN J. C. GAULIN³

¹Department of Psychology, University of California, Santa Barbara, CA 93106, USA. ²Department of Psychology, Yale University, New Haven, CT 06520, USA. ³Department of Anthropology, University of California, Santa Barbara, CA 93106, USA.

References

1. J. New, M. M. Krasnow, D. Truxaw, S. J. C. Gaulin, *Proc. Biol. Soc.* **274**, 2679 (2007).
2. D. F. Halpern, *Sex Differences in Cognitive Abilities* (Lawrence Erlbaum, Mahwah, NJ, ed. 3, 2000).
3. F. W. Marlowe, *Evol. Anthropol.* **14**, 54 (2005).
4. C. B. Stanford, in *Meat-Eating and Human Evolution*, C. B. Stanford, H. T. Bunn, Eds. (Oxford Univ. Press, Oxford, 2001), pp. 122–140.
5. U. Dogu, F. Erkip, *Environ. Behav.* **32**, 731 (2000).
6. M. Daly, M. Wilson, *Sex, Evolution and Behavior* (PWS, Belmont, CA, ed. 2, 1983).

The ABCs of Multiple Bonding

IN A RECENT PERSPECTIVE (6 APRIL, P. 61) discussing high bond orders in metal-metal bonding and in the silicon homolog of acetylene $\text{HC}\equiv\text{CH}$, where $\text{C} = \text{Si}$, F. Weinhold and C. R. Landis state that “the first stable $\text{Si}\equiv\text{Si}$ species was reported to adopt a planar, but nonlinear, trans-bent geometry,” citing the work by Sekiguchi *et al.* (1). They show structure **C** (Fig. 1) for the parent compound HSiSiH and explain the difference to the bonding situation in linear acetylene HCCH in terms of high preference for p-orbital character in the Si-Si σ bond (2). However, the compound HSiSiH was already synthesized in 1991 (3), 15 years earlier than Sekiguchi *et al.*, who isolated the substituted derivative RSiSiR

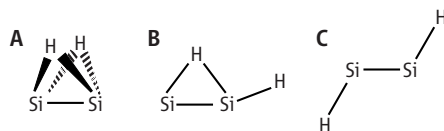


Fig. 1. Energetically lowest lying structures **A**, **B**, and **C** of HSiSiH . The lines indicate only the atomic connectivities but not the electron pairs of Lewis structures. Relative energies in kcal/mol are from (6).

where R is a bulky silyl group (1). Destombes and co-workers (3) showed that the equilibrium geometry of HSiSiH does not form structure **C**, but rather the doubly bridged structure **A** (Fig. 1) previously predicted by ab initio calculations (4). Photoexcitation of **A** yields another isomer of HSiSiH , identified as structure **B** (Fig. 1) (5). Quantum chemical calculations predict that structures **A** and **B** are lower in energy than structure **C** (6).

Thus, the silicon homolog of acetylene does not exhibit the Lewis-like bonding pattern discussed by Weinhold and Landis, who consider only the energetically high-lying form **C**. The trans-bent geometry of the RSiSiR compound of Sekiguchi *et al.* (1) comes from steric repulsion between the bulky R groups, which prevent the formation of the isomeric forms **A** and **B**. Sekiguchi's compound RSiSiR is chemically more stable than HSiSiH , which can only be isolated in a low-temperature plasma (3, 5), but this results not from more favorable Si-Si binding interactions but rather from the steric protection of the silicon atoms.

Why do the energetically lowest lying structures of the silicon homolog of acetylene exhibit the unusual hydrogen-bridged geometries **A** and **B** (Fig. 1)? We have analyzed the interactions between two EH fragments (where $\text{E} = \text{C}, \text{Si}, \text{Ge}, \text{Sn}, \text{Pb}$) in the electronic ground state and the first excited state (6). It is only the

combination of two EH fragments in the first excited state that leads to the standard Lewis-type structure with a linear arrangement $\text{HE}\equiv\text{EH}$, because each fragment has three unpaired electrons that yield a triple bond. In contrast, the interactions between EH fragments in the ground state—which has one electron pair, one empty orbital, and only one unpaired electron—explain the preference for **A** and **B** over **C**. Acetylene differs from its heavier homologs, because it takes much less energy to excite CH from the ground state to the excited state than for the heavier species EH. Only in the carbon compound does the stronger bonding in the linear form $\text{HC}\equiv\text{CH}$ compensate for the excitation energy of the fragments.

The experimental (1) and theoretical (6) studies do not agree with the conclusion of Weinhold and Landis that “[f]uture synthetic and computational explorations should be guided by closer attention to the maximally matched donor-acceptor interactions that lead to favorable Lewis-type bonding patterns.” Careful examination of the experimental and

theoretical findings for molecules that have high bond orders between metals clearly shows that the bonding in these compounds is not properly described by simple Lewis-type structures. Future work should use methods and ideas that are not confined to classical bonding models.

GERNOT FRENKING AND RALF TONNER

Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, D-35043 Marburg, Germany.

References

1. A. Sekiguchi, R. Kinjo, M. Ichinohe, *Science* **305**, 1755 (2004).
2. C. R. Landis, F. Weinhold, *J. Am. Chem. Soc.* **128**, 7335 (2006).
3. M. Bogey, H. Bolvin, C. Demuynck, J.-L. Destombes, *Phys. Rev. Lett.* **66**, 413 (1991).
4. H. Lischka, H. Köhler, *J. Am. Chem. Soc.* **105**, 6646 (1983).
5. M. Cordonnier, M. Bogey, C. Demuynck, J.-L. Destombes, *J. Chem. Phys.* **97**, 7984 (1992).
6. M. Lein, A. Krapp, G. Frenking, *J. Am. Chem. Soc.* **127**, 6290 (2005).

Response

THE STRUCTURAL ISOMERS AT ISSUE CONTAIN Si-Si bonds with either one bridging H (Cs point group symmetry), two bridging H's

CORRECTIONS AND CLARIFICATIONS

News Focus: “Greening the meeting” by B. Lester (5 October, p. 36). The analyses of the AGU annual conference mentioned on page 36 and the ESA conference on page 37 were done by Lawrence Plug and Borden Scott of Dalhousie University, not David Scott. The abstract of their poster was published in *Eos Trans. AGU Fall Meet. Suppl.* **84** (2003). Also, the credit for the bottom image on p. 38 should read “NCSA ACCESS (Arlington, VA)/WNT Consulting LLC Architecture and Innovative Technology Design.”

Cover Caption: (28 September, p. 1821). In the credit line, Andrea Ottesen's name was misspelled.

Editorial: “Playing climate change poker” by C. Challen (20 July, p. 295). The phrase “intended to reduce average global warming by 2°C” was meant to signify that average global warming be limited to 2°C, not that the current average be reduced by 2°C.

TECHNICAL COMMENT ABSTRACTS

COMMENT ON “A Vestige of Earth's Oldest Ophiolite”

Allen P. Nutman and Clark R. L. Friend

Furnes *et al.* (Reports, 23 March 2007, p. 1704) reported the identification of an ophiolite sequence within the ~3.8-billion-year-old Isua supracrustal belt. However, they did not acknowledge that the belt contains supracrustal rocks and mafic dikes of different ages, nor did they demonstrate that the proposed components of the ophiolite are coeval.

Full text at www.sciencemag.org/cgi/content/full/318/5851/746c

COMMENT ON “A Vestige of Earth's Oldest Ophiolite”

Warren B. Hamilton

The claim by Furnes *et al.* (Reports, 23 March 2007, p. 1704) that Greenland metavolcanic rocks require Paleoproterozoic seafloor spreading is incompatible with their own data. The purported sheeted dikes have the composition of pyroxenitic komatiite and could not have fed the adjacent ferroandesitic pillow lavas. Neither type has ophiolitic analogs, and both are likely ensialic.

Full text at www.sciencemag.org/cgi/content/full/318/5851/746d

RESPONSE TO COMMENTS ON “A Vestige of Earth's Oldest Ophiolite”

Harald Furnes, Maarten de Wit, Hubert Staudigel, Minik Rosing, Karlis Muehlenbachs

The comments by Nutman and Friend, and Hamilton, question our evidence for the presence of the Isua ophiolite. Their critical remarks are particularly directed at the veracity of our inferred sheeted-dike complex, the cogenicity of pillow lavas and dikes, and the nonexistence of modern equivalents. Here, we expand on our explanations in response to each of their comments to better justify our arguments and interpretation.

Full text at www.sciencemag.org/cgi/content/full/318/5851/746e

(C_{2v} point group symmetry), or no bridging H's (trans-bent, C_{2h} point group symmetry), all of which are local minima. These alternative structures were duly noted in Sekiguchi *et al.*'s original *Science* Report of the first Si≡Si triple bond (1), the accompanying Perspective (2), and in our own recent Research Article (3). However, in the context of our Perspective (6 April, p. 61), such H-bridged isomers are irrelevant: They do not contain Si≡Si triple bonds. Natural bond orbital (NBO) and natural resonance theory (NRT) analysis of the H-bridged isomers yields leading Lewis structures with Si-Si single-bonding in the C_{2v} isomer and double-bonding in

the Cs isomer (additional shared density in Si-H-Si interactions gives net Si-Si NRT bond orders of 1.8 and 2.5, respectively), in contrast to the triple bonding (NRT bond order of 2.9) in the cited C_{2h} isomer (4). Fuller discussion of these alternative isomers and the interesting electronic origins of the preference for H-bridged versus unbridged bonding motifs was precluded by considerations of length and relevance to the main Perspective topic.

Our Perspective sought to address the fundamental question, "How many bonds can be made between two atoms?" For silicon, Sekiguchi's macroscopic-scale synthesis and crystallographic characterization of the persistent RSiSiR (R=CH(SiMe₃)₂) first suggested that the answer was three. Computational analysis of trans-bent Si₂H₂, which bears strong geometric and electronic similarity to Sekiguchi's compound, supports this formulation. Ongoing synthetic studies to explore the upper limits of bonding between two atoms likely will feature bulky R substituents rather than H for two very different reasons. First, the steric protection provided by bulky attachments hinders alternate reaction pathways and eases isolation. Second,

such substituents disfavor bridging modes that necessarily lower the maximum achievable bond order. As stated in our Perspective, the highest achievable bond orders most likely will be realized by "interactions that lead to favorable Lewis-type bonding patterns," as clearly demonstrated in the theoretical logic that led to successful prediction (5, 6) of the Cr-Cr quintuple-bonding motif prior to its recent synthesis (7).

FRANK WEINHOLD AND CLARK R. LANDIS

Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, WI 53706, USA. E-mail: landis@chem.wisc.edu (C.R.L.); weinhold@chem.wisc.edu (F.W.)

References and Notes

1. A. Sekiguchi, R. Kinjo, M. Ichinohe, *Science* **305**, 1755 (2004).
2. R. West, *Science* **305**, 1724 (2004).
3. C. R. Landis, F. Weinhold, *J. Am. Chem. Soc.* **128**, 7335 (2006).
4. We note that these NRT bond orders are wholly consistent with the Si-Si electron pair sharing scheme presented by the Frenking group; M. Lein, A. Krapp, G. Frenking, *J. Am. Chem. Soc.* **127**, 6290 (2005).
5. F. Weinhold, C. R. Landis, *Valency and Bondin: A Natural Bond Orbital Donor-Acceptor Perspective* (Cambridge Univ. Press, Cambridge, 2005), pp. 555–559.
6. F. Weinhold, C. R. Landis, *Chem. Educ. Res. Practice Europe* **2**, 91 (2001).
7. T. Nguyen *et al.*, *Science* **310**, 844 (2005).

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

Science Classic

The complete
Science archive
1880–1996

Fully integrated with
Science Online
(1997–today)

Available to institutional
customers through a site license.
Contact ScienceClassic@aaas.org
for a quote.

Information: www.sciencemag.org/classic



Comment on "A Vestige of Earth's Oldest Ophiolite"

Allen P. Nutman^{1*} and Clark R. L. Friend²

Furnes *et al.* (Reports, 23 March 2007, p. 1704) reported the identification of an ophiolite sequence within the ~3.8-billion-year-old Isua supracrustal belt. However, they did not acknowledge that the belt contains supracrustal rocks and mafic dikes of different ages, nor did they demonstrate that the proposed components of the ophiolite are coeval.

Furnes *et al.* (1) reported that a sheeted-dike complex they identified within the ~3.8-billion-year-old Isua supracrustal belt (ISB) in Greenland provides the oldest evidence of oceanic crustal accretion by spreading. However, they did not alert readers that the ISB contains supracrustal rocks and mafic dikes of different ages (2, 3). They also failed to demonstrate that the proposed components for their ophiolite are coeval. These are important oversights, because a genuine ophiolite is a coeval assemblage of gabbros, sheeted dikes, and pillow lavas [e.g., (4)].

At Isua, the copious Paleoproterozoic Amlalik dyke swarms cut all Eoarchean rocks, including all components of the 3.81 to 3.63 Ga orthogneisses (Fig. 1) that envelop the ISB (5). Within the ISB, the Amlalik dikes are variably deformed and largely recrystallized into amphibolites (6). Remarkably, Furnes *et al.* (1) did not even mention that these dikes exist when discussing the origin of their "sheeted dikes." In the area covering localities 2 and 3 in (1), detailed mapping (Fig. 2) shows that there are numerous amphibolite dikes of differing thickness that are aligned subconcordantly to the lithological layering of the host volcano-sedimentary rocks. Dikes occur not only in the metavolcanic amphibolites as described in (1), but also in siliceous metasediments, ultramafic rocks, and the petrogenetically unrelated "boninitic" amphibolites to the west. As these dikes cut a wide range of unrelated lithologies, they cannot all represent a simple "sheeted dike" complex as proposed. Furthermore, in geochemical diagrams [Fig. 3 in (1)], the data presented show that their dikes are less evolved than the material they are supposed to feed. This is contrary to the suggestion that the pillows and sheeted dikes are related. Thus, Furnes *et al.* (1) need to show that they have distinguished dikes that are younger, unrelated intrusions such as (<3.5 Ga) Amlalik dikes and ones that might really form an earlier sheeted complex.

Furnes *et al.* stated that traversing northwards from their localities 3 to 1 entails passing stratigraphically upwards from sheeted dikes to pillows [figure 2A in (1)]. Thus, pillows at their locality 1 should be facing northward. Our photograph of the same pillows (Fig. 3) shows that they actually face southward, opposite to the sense that is required for their proposed simple stratigraphic relationship. Therefore the struc-

tural relationships cannot be as simple as suggested in (1).

Furnes *et al.* did not explicitly inform their readers that the Isua supracrustal belt contains fragments of both 3.7 and 3.8 Ga volcano-sedimentary sequences (2, 7). Thus, the metachert unit that crops out beside their localities 1 to 3 (Figs. 1 and 2) contains rare ~3.7 Ga volcano-sedimentary zircons (7), suggesting the maximum age of this package as ~3.7 Ga. However, in the southwest of the belt, where Furnes *et al.* (1) proposed that there are coeval ophiolitic gabbroic protoliths, amphibolites are cut by ~3.8-Ga tonalite sheets, giving their minimum age as ~3.8 Ga (2, 7, 8). Have Furnes *et al.* equated ~3.8 Ga metagabbros with ~3.7 Ga metavolcanic rocks and perhaps some still younger dikes? More exacting evidence is required to turn "Earth's oldest ophiolite" from an exciting proposition into an established fact.

References

1. H. Furnes, M. de Wit, H. Staudigel, M. Rosing, K. Muehlenbachs, *Science* **315**, 1704 (2007).

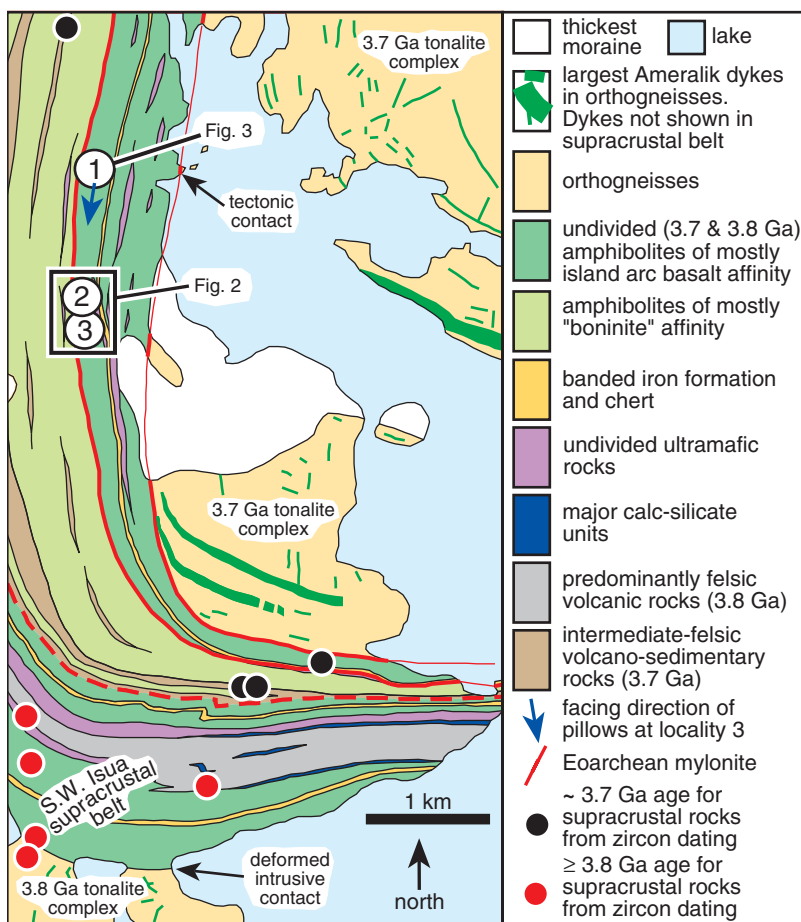


Fig. 1. Geological sketch map of part of the western end of the Isua supracrustal belt. Only the thickest, most continuous areas of cover moraine are shown. Mapping compiled from (6, 7). Zircon dating results constraining the ages of supracrustal rocks in the southwest to ≥ 3.8 Ga and in the northeast to ≤ 3.7 Ga are shown. The localities 1, 2, and 3 are from (1). The belt is partitioned by Eoarchean shear zones. The likely position of the break between the ~3.8 and 3.7 Ga sequences is presently known only within 200 m (7).

¹Institute of Geology, Chinese Academy of Geological Sciences, 26 Baiwanzhuang Road, Beijing 100037, China.

²45 Stanway Road, Headington, Oxford OX3 8HU, UK.

*To whom correspondence should be addressed. E-mail: nutman@bjshrimp.cn

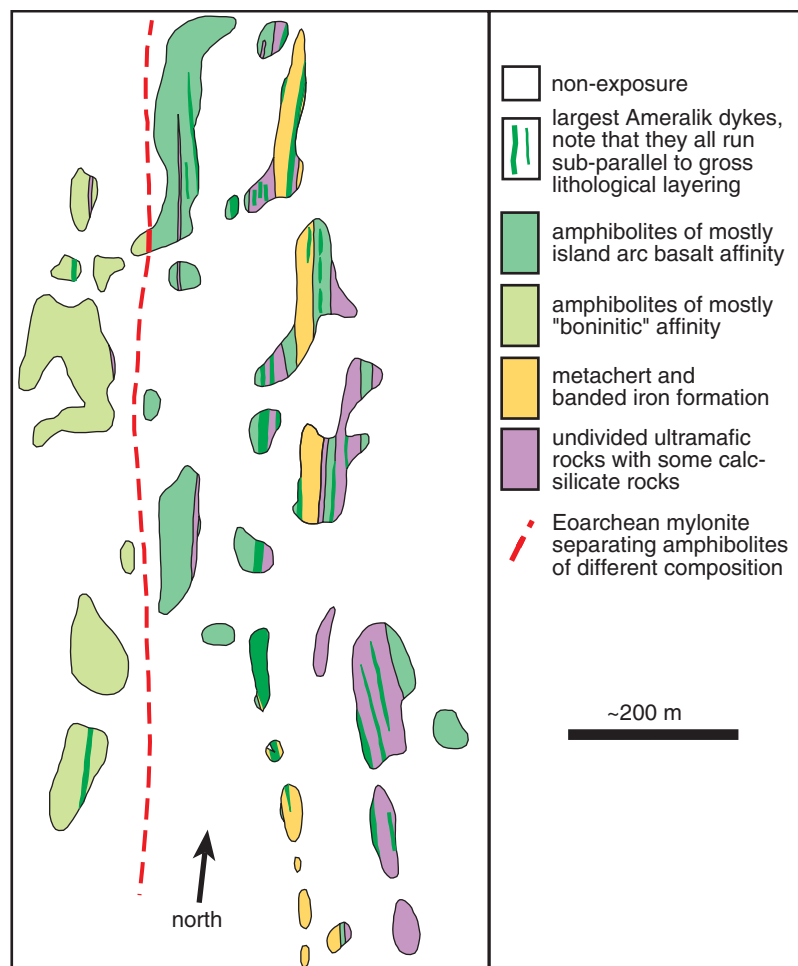


Fig. 2. About 1:10,000 scale mapping by A. P. Nutman in August 1980 of the area thought to cover the Furnes *et al.* localities 2 and 3, by using their low-resolution sketch map [figure 1C in (1)]. Numerous amphibolitized dikes cut all lithologies and are oriented subconcordant to the lithological layering.

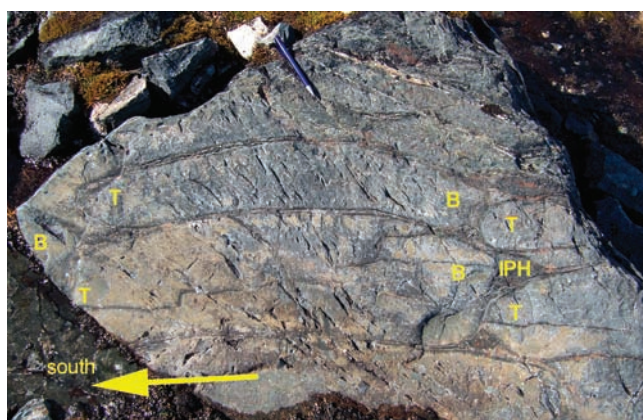


Fig. 3. Pillow lavas at Furnes *et al.* locality 1. The shape of the pillows indicates facing to the south (left of picture), and hence toward the proposed sheeted dikes. IPH is the interpillow hyaloclastite shown by Furnes *et al.*; T and B are the top and base, respectively, of some pillows. The pillows have been flattened orthogonal to their original orientation. [Photo by A. P. Nutman]

2. A. P. Nutman, V. C. Bennett, C. R. L. Friend, M. T. Rosing, *Chem. Geol.* **141**, 271 (1997).
3. A. P. Nutman, C. R. L. Friend, V. C. Bennett, *J. Geol. Soc. London* **161**, 421 (2004).
4. J. W. Shervais, *Geochem. Geophys. Geosyst.* **2**, paper 2000GC000080 (2001).
5. J. H. Allaart, in *The Early History of the Earth*, B. F. Windley, Ed. (Wiley, London, 1976), pp. 177–189.
6. A. P. Nutman, *Bull. Grøn. Unders.* **154**, 88 (1986).
7. A. P. Nutman, C. R. L. Friend, V. C. Bennett, *Tectonics*, **21**, article 5 (2002).
8. J. L. Crowley, *Precambrian Res.* **126**, 235 (2003).

23 April 2007; accepted 7 September 2007
10.1126/science.1144148

Comment on "A Vestige of Earth's Oldest Ophiolite"

Warren B. Hamilton

The claim by Furnes *et al.* (Reports, 23 March 2007, p. 1704) that Greenland metavolcanic rocks require Paleoproterozoic sea-floor spreading is incompatible with their own data. The purported sheeted dikes have the composition of pyroxenitic komatiite and could not have fed the adjacent ferroandesitic pillow lavas. Neither type has ophiolitic analogs, and both are likely ensialic.

When did plate tectonics (the rifting and aggregation of internally rigid plates of lithosphere, as oceanic lithosphere formed by sea-floor spreading between some of them and was subducted beneath others) begin? One bit of the voluminous evidence for a late start is the lack of sections of ancient oceanic lithosphere, ophiolites, abundant in Phanerozoic terrains and consisting of submarine mafic and ultramafic igneous rocks erupted directly upon residual mantle. Furnes *et al.* (1) espouse the contrary view that plate tectonics began very early and reported the discovery of a partial Paleoproterozoic ophiolite in the form of "cogenetic" metamorphosed and deformed sheeted dikes and pillow lavas of "intraoceanic island arc and mid-ocean ridge basalt (MORB) affinities." However, their data neither support these inferences nor provide the claimed "compelling structural evidence of horizontal extension by dike injection at a spreading ridge."

Neither the Isua pillow lavas nor the purported dikes fit into the broad array of modern island-arc and seafloor basalts for which Furnes *et al.* use the term "ophiolitic." The nine similar analyses of "dikes" [table S1 in (1)] are ultramafic, with $\text{CaO} > \text{Al}_2\text{O}_3$, 48 to 51% SiO_2 , ~15 to 20% MgO , ~12 to 13% Fe_2O_3^* (total iron as Fe_2O_3), and ~1% Na_2O . This is the bulk composition of pyroxenitic komatiite, an Archean lava type (2), probably ensialic elsewhere, common at Isua (3, 4) (primary textures have been obliterated by metamorphism). The adjacent pillow lavas cannot have been erupted from the dikes: The five subuniform analyses of lavas [table S1 in (1)] have $\text{Al}_2\text{O}_3 \gg \text{CaO}$, 54 to 57% SiO_2 , ~5% MgO , ~12% Fe_2O_3^* , and ~4% Na_2O and have the composition of mafic ferroandesite, an

Archean rock type that is ensialic elsewhere. The near uniformity of each type makes gross metasomatic change of relatively immobile elements implausible. Neither rock type has common modern equivalents, in or out of ophiolites. Sheeted dikes in Phanerozoic ophiolites are, by contrast, compositionally similar to, or moderately less feldspathic than, associated pillow lavas [e.g., (5, 6)], and are not ultramafic.

These Isua pillow lavas do reasonably match, in major elements and minor-element ratios, widespread rocks in the Australian kilometers-thick Neoproterozoic Fortescue sheet of mostly mafic flood lavas, which overlies 150,000 km^2 of older Archean continental crust (7). Fortescue rocks are unlike any modern continental flood basalts, and Ti/Y/Zr discriminants similar to those used by Furnes *et al.* misclassify them as mostly oceanic-arc basalts (7), thus invalidating conjecture that numerology can define tectonic settings of Archean rocks in the absence of geologic information. Isua stratigraphy is undefined, but the lower parts of Archean supracrustal sections commonly have, like little-deformed Fortescue, subregional sheet stratigraphy where constrained by good mapping and dating.

Archean igneous rocks contrast strikingly with modern ones in both bulk compositions and lithologic and stratigraphic assemblages, precluding the direct analogies with modern tectonic settings possible in Phanerozoic terrains. Investigators [e.g., (1)] who nevertheless seek uniformitarian plate-tectonic explanations commonly appeal to selected minor-element ratios as discriminants, with the assumption that those ratios somehow define tectonic settings. That the discriminants misclassify not only Fortescue, but also many modern assemblages of known settings, and that none of the structural and associational predictions implicit in derivative plate-tectonic assignments are fulfilled in Archean terrains, has

not deterred the widespread practice by geochemists. The supposed suprasubduction Isua "boninites" (1, 8) are defined only with trace-element rationales, and their basaltic to high-Mg-basaltic, non-boninitic major-element compositions (e.g., low SiO_2) are ascribed to gross additive and subtractive metasomatism (8).

The Isua sheeted dikes inferred by Furnes *et al.* from a small area are severely flattened decimeters-thick layers of relatively massive rock, separated irregularly by more schistose rock assumed to represent chilled margins and volcanic rocks. If the layers indeed are derived from dikes, they indicate crustal extension but not its setting, because sheeted dikes occur locally in non-oceanic settings; or the layering may be inherited from planar decimeter- to meters-thick internal facies of close-following pyroxenitic-komatiite lavas (9) at a locality of simple triaxial strain and favorable preservation.

Archean greenstone belts commonly have, like Isua, thick sections dominated by mafic and ultramafic lavas. Only felsic basement, mostly tonalitic, trondhjemitic, and granodioritic gneisses, has ever been found depositionally beneath them, and clastic sediments derived from that basement commonly intervene between it and the volcanic sections where not cut out by younger magmatism [e.g., (10)]. No Archean supracrustal rocks are proved ensialic. Furnes *et al.* did not consider non-plate-tectonic alternatives incorporating delamination (11).

References and Notes

1. H. Furnes, M. de Wit, H. Staudigel, M. Rosing, K. Muehlenbachs, *Science* **315**, 1704 (2007).
2. A. Riganti, A. H. Wilson, *Lithos* **34**, 235 (1995).
3. A. Polat, A. W. Hofmann, *Precambrian Res.* **126**, 197 (2003).
4. Twelve of the 13 "least altered" samples from the "inner arc sequence" in (3) [which includes the localities of Furnes *et al.* (1)] have the distinctive bulk composition of pyroxenitic komatiite. Polat *et al.* (8) assumed the rocks to be island-arc picrites, a quite different uncommon modern rock type.
5. H. Staudigel *et al.*, *Geochem. Geophys. Geosyst.* **1**, Paper 1999GC1 (1999).
6. M. A. Stewart *et al.*, *J. Geophys. Res.* **108** (B4), Paper EPM-2 (2001).
7. A. M. Thorne, A. F. Trendall, *Geol. Surv. W. Australia Bull.* **144**, 249 (2001).
8. A. Polat, A. W. Hofmann, M. T. Rosing, *Chem. Geol.* **184**, 231 (2002).
9. A. H. Wilson, J. A. Versfeld, D. R. Hunter, *Contribs. Mineral. Petrol.* **101**, 301 (1989).
10. W. Bleeker, *Geol. Soc. London Spec. Publ.* **199**, 151 (2002).
11. W. B. Hamilton, *Geol. Soc. Am. Mem.* **200**, 233 (2007).

10 May 2007; accepted 1 October 2007
10.1126/science.1144931

Department of Geophysics, Colorado School of Mines, Golden, CO 80401, USA. E-mail: whamilton@mines.edu

Response to Comments on “A Vestige of Earth’s Oldest Ophiolite”

Harald Furnes,^{1*} Maarten de Wit,² Hubert Staudigel,³ Minik Rosing,⁴ Karlis Muehlenbachs⁵

The comments by Nutman and Friend, and Hamilton, question our evidence for the presence of the Isua ophiolite. Their critical remarks are particularly directed at the veracity of our inferred sheeted-dike complex, the cogenicity of pillow lavas and dikes, and the nonexistence of modern equivalents. Here, we expand on our explanations in response to each of their comments to better justify our arguments and interpretation.

Our study (1) emphasized that the components of the Isua ophiolite are tectonically dismembered and that its reconstruction is therefore not a trivial task, as compared with many younger (Phanerozoic) ophiolites. Hence, we welcome the comments by Nutman and Friend (2) and Hamilton (3) as an opportunity to provide further details, to clarify the re-construction of the Isua ophiolite, and to respond to the following issues raised in their comments: (i) the sheeted dikes, (ii) the geochemical composition of the dikes and the pillows, (iii) the facing (stratigraphic) relationships between the dikes and pillows, (iv) geochronological constraints, (v) comparison with modern equivalents, and (vi) implications for Archean plate tectonics.

Nutman and Friend (2) state that “At Isua, copious Paleoproterozoic Ameralik dike swarms cut all Eoarchean rocks, including all components of the 3.8 to 3.6 Ga orthogneisses that envelop the Isua supracrustal belt” and imply that we may have misidentified the Ameralik dikes as (part of) an ophiolitic sheeted-dike complex. The sheeted dikes from our study are clearly distinguishable from the Ameralik dikes. The latter occur generally as thick tabular bodies with ophitic texture and plagioclase megacrysts. They are largely undeformed and, in the area of question, lack cleavage or schistosity. They cross-cut metacherts, pillow lavas, ultramafic units, and the garbenschiefer rocks (including the boninitic metabasalts), which flank the sheeted-dike complex. The sheeted-dike complex in contrast forms tabular mappable units which contain thin screens of pillow lava, volcanic breccias [see figure 2C in (1)], and occasional plagiogranite, and comprise

highly deformed and foliated amphibolites that are in places crosscut, often orthogonally, by undeformed Ameralik dikes. The Ameralik dikes hence postdate deformation of the ophiolitic sequence and, in the area of question, are easily distinguished from the highly deformed ophiolitic dikes. Although the “low-resolution” sketch maps (1, 2), both modified directly from (4), are not at appropriate scales to illustrate these details, we are confident that we have correctly differentiated the younger Ameralik dikes from the amphibolites interpreted by us to be sheeted dikes.

Our chemical data [table S1 in (1)] show that the dikes of the sheeted complex are less evolved than the pillow lavas. Nutman and Friend (2) and Hamilton (3) state that the pillow lavas cannot be derived from the dikes because they are geochemically too different; hence, they question the inferred cogenicity. We disagree with this assessment, because immobile trace elements (e.g., Ti-V and Ti-Zr) covary along common fractionation trends, and the ratios of incompatible elements (e.g., Zr/Y) are very similar between the dikes and the pillows. Moreover, the dikes must necessarily be closer to the magma source than the pillow lavas that they fed, and we argue that the magma underwent fractional crystallization en route to the surface. Although the major element variations of the Isua rocks are likely complicated by metamorphic overprints, they can largely be explained by fractional crystallization of Mg-rich olivine and Ca-rich pyroxene leading to the chemical differences between the dikes and pillow lavas seen. Such chemical differences are common in most volcanic systems.

Hamilton (3) uncritically uses the major element composition of the dikes to classify them as “pyroxenitic komatiites.” Such classifications should not, however, be made on the basis of elements that may have been mobile during metamorphic conditions. For example, the process of spilitization—which causes local redistribution of principally Ca and Mg, together with a gain in Si, Na, and H₂O and often a loss of Ca, K, Sr, Rb, and Ba—is a well-known process [e.g., (5)] that has misled many geochemists. Spilitization is commonplace in Archean volcanic rocks,

and we are currently documenting such chemical changes in the little-deformed, low-grade metamorphic lavas from the Mesoproterozoic Barberton Greenstone Belt in South Africa [the type location of (pyroxenitic) komatiites], from which we have geochemical data for more than 500 samples. Large chemical changes may have occurred in the Isua complex and, hence, we reject this classification. Hamilton (3) further draws attention to the similarity between the little-deformed basalts of the ~2.7 Ga Fortesque Group of the Pilbara Craton (Western Australia) and the geochemistry of the Isua dikes and lavas, and to the former having been misclassified on the basis of Ti/Y/Zr discriminant diagrams (6). We do not see the relevance of these observations extrapolated from the Pilbara Craton to the discussion about Isua.

Nutman and Friend (2) point out that the facing (younging) direction of the pillows is opposite to the transition from the dike to pillow facies. This younging relationship, which has been determined in one outcrop by them and by us, has little relevance because the ophiolite units cannot be traced continuously and the pillows and dikes are strongly flattened and stretched, and thus tectonically rotated, a feature which is also common in Phanerozoic ophiolites [e.g., (7)]. Furthermore, the supracrustal belt has experienced several episodes of deformation and folding [e.g., (4)]. In such a setting, the attitude of dikes, bedding, and the younging direction of pillows can attribute little, if any, relevance to their original disposition. It is thus not possible to use a local facing direction as a regional reconstruction tool without a more complete understanding of the structural geology.

Nutman and Friend (2) claim that the existing geochronological constraints do not support a cogenetic relationship between the components of the Isua ophiolite that we have identified. There are, however, no radiometric dates on the individual mafic-ultramafic components (i.e., pillow lavas, sheeted dikes, gabbro, and ultramafics) that constitute the Isua ophiolite to corroborate this. As we pointed out in (1), current whole-rock Sm-Nd age data from metasediments and enclosing garbenschiefer yield an age of $3,779 \pm 81$ Ma (8), and the pillow lavas and metagabbro an age of $3,777 \pm 41$ Ma (9). Indeed, detailed chronostratigraphy within any of the Isua tectonic packages is still lacking, and several geochronological aspects of the Isua belt remain unresolved. Lack of geochronological data cannot therefore be the basis for discounting cogenicity.

Hamilton (3) states that we stretch the term “ophiolite” both in terms of construction and petrological development and also questions whether modern equivalents for Archean oceanic crust, and hence the Isua ophiolite, exist. If the term “ophiolite” is to be restricted to the typical Phanerozoic “Penrose-type” ophiolite only, we might agree. There exists, however, a much broader classification scheme for ophiolites that includes

¹Centre for Geobiology and Department of Earth Science, University of Bergen, Bergen, Norway. ²AEON-Africa Earth Observatory Network, and Department of Geological Sciences, University of Cape Town, Rondebosch 7700, South Africa. ³Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0225, USA. ⁴Nordic Centre for Earth Evolution (NordCEE), and Geological Museum, University of Copenhagen, 1350 København K, Denmark. ⁵Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta T6G 2E3, Canada.

*To whom correspondence should be addressed. E-mail: harald.furnes@geo.uib.no

a wide variation in terms of lithological construction and petrological evolution, depending on their environment of formation and stage of evolution (10). For example, the Caribbean-type ophiolites that represent the oceanic crustal assemblages common in Large Igneous Province (LIP) [e.g., (10, 11)] distinctly differ from common Penrose-type ophiolites, having thick volcanic piles (exceeding 5 km) that include substantial volumes of ultramafic rocks [e.g., (12, 13)]. It is also pertinent to mention that the Tortuga ophiolite complex, a Chilean-type ophiolite (10), contains high-Mg basalt dikes cutting the deeper diabase and gabbro units below the sheeted-dike complex (e.g., 14).

Hamilton's views on the Archean are well known [e.g., (15)], and we do not wish to repeat the discussion of Hamilton (16) and counterarguments by de Wit (17). Our Isua field observations and new geochemical data put this debate on a new footing, and we are striving to allow the rocks themselves to provide the answers to how the Earth worked at 3.8 Ga. Our physical evi-

dence for a spreading center [figure 2, C to F, in (1)] and geochemical evidence for subduction-related magmatic activity from Isua (3, 18, 19) offer evidence that is consistent with Archean plate tectonics, but much remains to be done to understand its full importance.

In summary, many of the comments by Nutman and Friend (2) and Hamilton (3) are pertinent, focusing on uncertainties in the complex geological history of the Isua supracrustal belt, including its inferred oceanic crustal component (1). Although we maintain our interpretation that these rocks represent part of a dismembered ophiolite, we agree that there are still many uncertainties left, and we reiterate from our paper (1) that there is a need for more in-depth work, based on detailed field observations.

References

1. H. Furnes, M. de Wit, H. Staudigel, M. Rosing, K. Muehlenbachs, *Science* **315**, 1704 (2007).
2. A. P. Nutman, C. R. L. Friend, *Science* **318**, 746 (2007); www.sciencemag.org/cgi/content/full/318/5851/746c.
3. W. B. Hamilton, *Science* **318**, 746 (2007); www.sciencemag.org/cgi/content/full/318/5851/746d.
4. A. P. Nutman, *Greenland Geol. Surv. Bull.* **154**, 80 (1986).
5. P. A. Floyd, *J. Petrol.* **14**, 522 (1976).
6. M. A. Thorne, A. F. Trendall, *Geol. Surv. W. Australia Bull.* **144**, 249 (2001).
7. H. Furnes et al., *Geochem. Geophys. Geosyst.* **4**, 1088 (2003).
8. M. T. Rosing, *Science* **283**, 674 (1999).
9. M. Boyet et al., *Earth Planet. Sci. Lett.* **214**, 427 (2003).
10. Y. Dilek, *Geol. Soc. Am. Spec. Pap.* **373**, 1 (2003).
11. M. F. Coffin, O. Eldholm, *Geol. Soc. Am. Spec. Pap.* **352**, 59 (2001).
12. G. Th. Klaver, *G. U. A. Pap. Geol. Ser.* 1, No. 27 (1987).
13. A. C. Kerr et al., *Contrib. Mineral. Petrol.* **124**, 29 (1996).
14. D. Elthon, *Nature* **278**, 514 (1979).
15. W. B. Hamilton, *Geol. Soc. Am. Today* **13**, 4 (2003).
16. W. B. Hamilton, *Precambrian Res.* **91**, 143 (1998).
17. M. de Wit, *Precambrian Res.* **91**, 181 (1998).
18. A. Polat, A. W. Hofman, *Precambrian Res.* **126**, 197 (2003).
19. A. Polat, A. W. Hofman, M. T. Rosing, *Chem. Geol.* **184**, 231 (2002).

21 May 2007; accepted 2 October 2007
10.1126/science.1144231

FISHERIES

Tales of a Small, But Crucial, Fish

Daniel Pauly

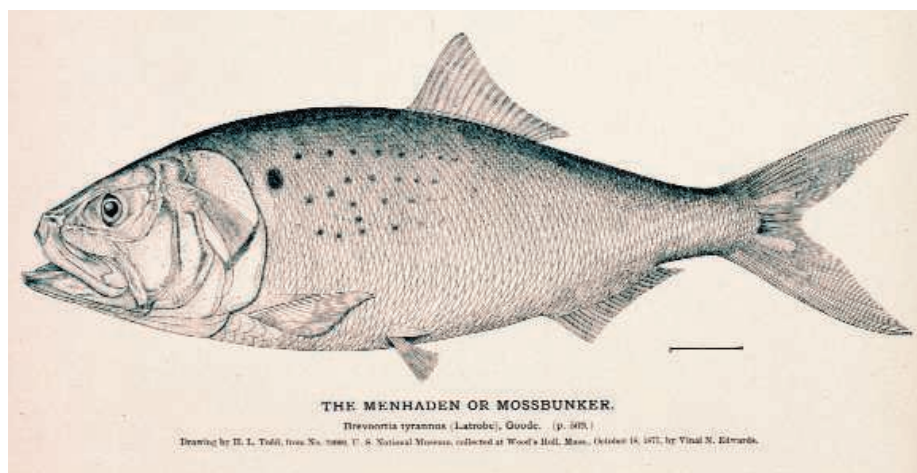
At first sight, *The Most Important Fish in the Sea* appears to belong to that lineage which earlier gave us accounts of cod [the fish that changed the world (1)], American shad [the founding fish (2)], and Patagonian toothfish, also known as Chilean seabass [the perfect fish (3)]. These fish were all important because of their impacts, past or present, on people's diets. But Atlantic menhaden (*Brevoortia tyrannus*) is really different from cod, shad, and toothfish, because it is a fish that we do not eat and likely never will. Rather, it is eaten by the fishes we like to catch and eat. Thus, the conflicts about and around this fish are different from the conflicts about others, where different people (the French versus the English, the line fishers versus the ones using trawls, etc.) competed for access to wholesome food fish.

With menhaden—an oily, bony, small, and reputedly ill-tasting representative of the herring family—the conflicts have been about the uses, direct or indirect, to which this fish was to be put. Traditionally, menhaden was used as fertilizer, and the book has an interesting section on how American Indians planted each corn plant with one fish. This role is at the origin of the fish's most common name, *munnowhatteaúg*, which means “that which manures” in the Algonquian language of the Narragansetts. On the other hand, the larger fish exploited by the early European settlers along the U.S. East Coast, and the marine mammals also abundant along that same coast, all fed on menhaden. This fish, and the microscopic algae it feeds on, formed the base and understory, respectively, of most coastal food webs, from New England to Florida and particularly in Chesapeake Bay.

With the invention of fish meal and its use for feeding chicken and livestock, the industrial fishery for menhaden increased tremendously, especially after World War II. Menhaden, which earlier had formed immense schools, immune to the frenzied hordes of predatory fishes surrounding them, became scarce. Their reduced numbers began to

affect the upper part of the food webs, threatening to drag all, prey and predators, into the maw of the reduction plants, which for a time mushroomed along the coast.

Conflict had always simmered between the



“A small, unappetizing fish.”

fishers exploiting larger fish, who wanted menhaden to “fulfill their natural role” (to be eaten by a large fish), and the reduction fishery (which employed spotter planes and purse seiners to save them from such cruel death). The debate intensified in the 1970s, when it

was joined by recreational anglers, whose target species (especially striped bass) depended on menhaden—despite assertions to the contrary by spokespersons (some, elected officials) of the reduction fishery.

This fisheries conflict was one of the first that pitted those interested in a single-species approach, hitherto dominant, against advocates of what is now called “ecosystem-based fishery management” (4). The arguments of both sides are still with us, even though (predictably) the bloated reduction fishery, along with the single stocks on which it depended, largely collapsed. All that is left in the mid-Atlantic region is a small stock of menhaden huddling in Chesapeake Bay and a single firm—the latter a distillate of everything that can be wrong with industrial fisheries (in particular, remote, but well-connected, corporate owners). Further north, off New Jersey and beyond, the now-protected menhaden are coming back.

In the Gulf of Mexico, the related Gulf menhaden (*Brevoortia patronus*) still supports an extensive fishery, generating conflicts that trail those along the Atlantic coast by one or two decades.

In the meantime, ecologists discovered that menhaden, given their feeding habits, were fulfilling another ecological role: keeping algal blooms in check. Although the role of oysters in cleaning up coastal waters was always understood, that of menhaden and

related filter-feeders was not. Now, people realize that it would be handy to have large schools of menhaden acting as giant vacuum cleaners in Chesapeake Bay and other coastal bodies currently choked by algal blooms triggered by farm runoffs.

The Most Important Fish in the Sea, which tells and thoroughly documents these stories, could be seen as yet another helpless commentary on the way we are trashing our oceans. But it is an optimistic book. It deals with a resilient little thing that, unlike larger, longer-lived species such as cod, readily bounces back if given the chance. The role of menhaden in coastal ecosystems is now well understood, making single-species arguments impossible to maintain. And the sole corporation that still fishes Atlantic menhaden for reduction does not have a monopoly in supplying fish meal and fish oil to the market. Indeed, it appears to be able to maintain its fleet only because of the welfare (subsidies) it gets. Perhaps this story will have a positive ending; H. Bruce Franklin's fascinating account makes us look forward to that.

References

1. M. Kurlansky, *Cod: A Biography of the Fish That Changed the World* (Walker, New York, 1997).
2. J. McPhee, *The Founding Fish* (Farrar, Straus, and Giroux, New York, 2002).
3. B. Knecht, *Hooked: Pirates, Poaching, and the Perfect Fish* (Rodale, New York, 2006).

The Most Important Fish in the Sea Menhaden and America

by H. Bruce Franklin

Island Press, Washington, DC, 2007. 278 pp. \$25.
ISBN 9781597261241.

The reviewer is at the Fisheries Centre, Aquatic Ecosystems Research Laboratory, 2202 Main Mall, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. E-mail: d.pauly@fisheries.ubc.ca

4. E. K. Pikitch et al., *Science* **305**, 246 (2004).
5. G. B. Goode et al., *The Fisheries and Fishery Industries of the United States* (Government Printing Office, Washington, DC, 1884–1887).

10.1126/science.1147800

FILM: SPACE

Reminiscences of the Moon Trips

Jay M. Pasachoff

As the 50th anniversary of the launch of Sputnik passed, a documentary film about the Apollo program gained widespread circulation. With its pedigree of “Ron Howard Presents,” and a cast of 8 of the 12 astronauts who landed on, and two others who circled, the Moon, *In the Shadow of the Moon* illuminates the heady era of lunar landings from a point of view nearly 40 years on.

Of course, it is remarkable that we—and in 1969 humanity took credit for the Moon landing—sent people off Earth during this past century. Sputnik or Apollo may be what is remembered from our previous millennium by people in the 30th or 40th centuries. It is also remarkable that we stopped going to the Moon 35 years ago.

The film, directed by David Sington (who has produced and directed science programs for television), is narrated by the astronauts, who are shown in extreme closeups. To at least some of us who remember them as young men with the Right Stuff (the Tom Wolfe title, to which some of them refer), it is a shock to see their white-haired heads on the screen. But after all, Buzz Aldrin, Neil Armstrong, and Michael Collins were all born in 1930, making them about 77. To today’s college students, the film remains one of derring-do. It tells the story leading up to and including the Apollo 11 mission, then deals cursorily with the missions following that first lunar landing. There is little about the Russian role in the space race. The movie incorporates footage that its researchers found in NASA’s Houston vaults.

The reviewer, co-author of *The Cosmos: Astronomy in the New Millennium*, is at Hopkins Observatory, Williams College, Williamstown, MA 01267, USA. E-mail: jay.m.pasachoff@williams.edu

In the Shadow of the Moon

David Sington, Director

THINKFilm, New York, in association with Discovery Films, 2007. 100 minutes. www.intheshadowofthemoon.com

The clear images reflect the restoration of the original films (no simulations or recreations are used).

We hear Alan Bean of Apollo 12, one of the more loquacious astronauts featured, describe how disbelieving he was when told on the phone that the Apollo 1 crew was “lost”: He first advised his caller to look for them in the beach house, before realizing the deadly consequences of the fire. We learn that Gus Grissom had been worried about the condition of the wiring in the 100%-oxygen atmosphere, but “I can’t say anything about it or they’ll fire me.” Perhaps it was the fire scene—and a glimpse of cigars lit in the Houston control room after Apollo 11’s landing—that brought the film its PG rating, for “mild language, brief violent images, and incidental smoking.”

One intriguing black-and-white sequence records the appearance of Mr. and Mrs. Armstrong, Neil’s parents, on the game show “I’ve Got a Secret.” Nobody guesses theirs: that their “son was made an astronaut today.” The interviewer then asked Mrs. Armstrong how she would feel if her son were chosen to land on the Moon.

Neil Armstrong, famously reclusive, chose not to appear in the film, though his name comes up often. Aldrin says that Armstrong’s cool manner was admirable, with his “One small step for man.” Aldrin admits that had he stepped out of the lander first, he might not have been able to refrain from shouting something like “Yahoo, man, I’m here.” He also describes how Armstrong decided to “go long” when a boulder field was under the lander and notes that “it was a little iffy right there at the very end.”

Michael Collins talks about how he felt about orbiting the Moon without getting a chance to land, pointing out that he was glad to have been on the crew of the first manned landing. He wasn’t lonely when on the far side of the Moon by himself, though he was aware that there were two people on the surface on the other side of the Moon and beyond them 3 billion on Earth, while “over here, there’s me plus... god only knows what.” He noted of Earth: “How fragile it appeared.” On his return, he

remarked, “Nice ocean you’ve got here, planet Earth.” And Bean recalls, “Since that time, I have not complained about the weather one single time. I’m just glad there is weather.”

The strangest piece of historic footage is Richard Nixon beginning a speech to the nation announcing the failure of Apollo 11 and loss of its crew. Fortunately, that speech never had to be delivered. Earlier, we were shown John Kennedy announcing the goal of bringing men to the Moon and back safely to Earth by the end of the decade. But the omission of Lyndon Johnson from the movie (we only see him sitting behind Kennedy during the latter’s Senate speech) seems very strange, because Johnson played major roles by suggesting the Moon landings to Kennedy and then by carrying through.

I was left with a sour taste by the film’s treatment of religion, near its end. Gene Cernan talks about a general creator “that stands above the religions that govern our lives,” and then Charlie Duke tells about finding Jesus. How



Eagle on the Moon. The Apollo 11 Landing Module, July 1969.

about the other 10 Moon-landing astronauts? Did any lose religion or decide that religion was not a particular part of his voyage to the Moon? Earlier, Jim Lovell, who read from Genesis on Apollo 8, responding to a letter from an atheist who wrote “that was inappropriate,” answers, “Maybe it was; I don’t know.”

All the same, Sington offers a moving tribute to “a time when we made bold moves.” The film’s final credits wonderfully proclaim, “This film was shot entirely on location on

A TASTE OF THE GONZO SCIENTIST

A Linnean Feast

If you haven't heard the news yet, this year marks the 300th birthday of Carl Linnaeus. I can report that his fellow Swedes are keenly aware. On a recent visit to Uppsala, I took part in a Linnaeus celebration organized by a group of Swedish scholars and chefs. It was called *Culina Mutata*, the changing kitchen, and it was the most unusual birthday party I have ever attended.

Linnaeus is best known as the inventor of modern taxonomy, the system of Latin names that divide all organisms into species based on shared traits. The innovation helped pave the way for Darwinism a century later. But like a rock star who is only remembered for that one catchy song, Linnaeus's other contributions are appreciated only by the groupies.

Less well-known is his passion for food. Toward the end of an illustrious career, Linnaeus laid the foundation for modern food science. Many of the questions he posed remain at the forefront of food science research today. Which foods are necessary for development, and which should be avoided? To what extent can diet promote or inhibit disease? Is there an ideal diet for each person, or indeed for each genome?

So for three days, a diverse medley of academics—biochemists, historians, agricultural scientists, psychologists—expounded on the science of food, both in the time of Linnaeus and today. The lectures were punctuated by a series of 18th-century meals, starting with a peasant's lunch and ending with a royal banquet. And for entertain-

ment? The conference organizers performed a play—in handmade period costumes, no less—about one extraordinary day in the life of Linnaeus. We were then exposed to the dance craze of the time, the minuet. (Imagine *Saturday Night Fever*, but with Mozart.)

If you missed the party, don't worry. The year is not yet done. Why not throw a Linnaeus 300th birthday party of your own? Use the following protocol to prepare a winter feast in true Linnaean style.

[The recipes (below) are from Gunnar Broberg and Gunilla Lindell's *Till livs med Linné* (Atlantis, Stockholm, 2007).]

First, find a fish. The winter months are best for the ruthlessly predatory but delicious pike (*Esox lucius*). They lurk beneath the lake ice. For a dinner party of four, catch two pike, if you can.

While your lines dangle, head into the forest to find a bird. An ideal quarry is the capercaillie (*Tetrao urogallus*),

also known as the wood grouse. But beware, they are intelligent and agile creatures. You will need a widely dispersing shotgun, if not an automatic weapon. Check what local laws apply.

On your way back to the lake with birds in hand, march to the swamp to gather cranberries (*Oxycoccus palustris*). You will find them in low depressions beneath the snow. If your feet get wet, build a fire. Better to ruin a dinner than lose a toe to the frost.

—JOHN BOHANNON
10.1126/science.1152652



For a full account see
www.sciencemag.org/sciext/gonzoscientist

Fried capercaillie

1 capercaillie
1 thin slice of lard
30 g fresh butter
salt and white pepper
Gravy:
45 g butter
20 g flour
350 ml cup cream
15 g tablespoon currant jelly
salt
white pepper

Pluck and gut the bird. Coat it with alcohol and set alight to singe the feather stubs. Rinse with water. Lightly coat the inside with salt. Bind the legs and fasten the sheet of lard over the breast with skewers. Sprinkle generously with salt and pepper. Put the bird on a spit and fry it with a blazing fire. Gather the fat as it falls, basting the bird frequently to prevent drying. (Do not underestimate this job. A kitchenboy is required.) For the

gravy, gather the excess fat and stir in flour before adding the cream and jelly. Cook in a saucepan for some minutes until it tastes right. If you managed to find cranberries, they would make a nice accompanying jelly.

Ice pike

2 pike, about 1.5 kg each
60 ml melted fresh butter
30 g grated horseradish
salt

Gut and rinse the fish but do not scale them. Place in boiling salted water until tender and serve with melted butter and grated horseradish. The dressing can be enlivened with a touch of vinegar and sugar. The fish should be presented first, along with drinking water, beer, barley bread, and the tale of how you caught them.

Winter cabbage

1 head red cabbage
3 apples
80 g goosefat (or butter)
1 onion
1 rind of pork
350 ml cup water
30 ml molasses or honey
15 ml vinegar
10 g salt
2.5 g white pepper

Melt the fat and cook shredded cabbage and sliced onion until lightly brown. Add peeled, sliced apples, the pork rind and some water. Simmer covered until the cabbage is tender. Add more water occasionally. Season carefully. The dish should have a mild and pleasant sour-sweet taste. Allow 2 hours for this.

Recipes adapted from Gunnar Broberg and Gunilla Lindell's *Till livs med Linné* (Atlantis, Stockholm, 2007).

Egg cheesecake with raisins

5 eggs
475 ml milk
115 g raisins
2.5 g teaspoon salt
honey to taste
5 ml teaspoon lemon juice

Rinse the raisins, dry them, and place in a cake pan. Meanwhile, thoroughly mix eggs and milk, followed by honey, salt, and lemon juice. Heat slowly. Remove from heat when it starts to curdle and beat slowly for a few minutes. Pour the mixture over the raisins and place the pan in a cool place. When the cake has thickened, serve with blueberry jam. You will be sleepy after this. Mind the candles. Do not burn down the house.

DISASTER MANAGEMENT

Confronting Disaster Losses

Laurens M. Bouwer,^{1*} Ryan P. Crompton,² Eberhard Faust,³ Peter Höppe,³ Roger A. Pielke Jr.⁴

Global costs of weather-related disasters have increased from an annual average U.S.\$8.9 billion (1977–1986) to U.S.\$45.1 billion (1997–2006). In coming decades, the number of people at risk from extremes will very likely grow, and extreme weather will likely increase (1). To date, societal change and economic development are mainly responsible for increasing losses. After adjusting for societal changes, loss time series reflect the climatological record (2).

By 2015, loss potentials among the world's 10 largest cities, most of which are in developing countries, are projected to increase from 22% (Tokyo) to 88% (Shanghai, Jakarta) (3). A repeat of the July 2005 floods in Mumbai (see figure, right) in 2015 could cause 80% higher losses and affect 20% more people, independent of climate change. Greenhouse gas emission reductions are of central importance, but they cannot decrease hazard risk for decades. In this context, we offer three recommendations for decision-makers.

Improve data collection. With few exceptions, records of disaster losses are of poor quality, inhomogeneous, and collected using a wide range of methods for different purposes, making research extremely challenging. Improved data could be used to evaluate disaster policies, estimate return periods, identify factors that drive loss trends and could potentially offer the prospect of an early-warning system for changes in the earth-climate system. Currently, the most comprehensive loss databases are held by insurance companies and are not publicly available. An open-source, peer-reviewed database would enable the scientific community to study worldwide disasters.

Expand the role of disaster risk reduction in adaptation. The cost-benefit ratio of disaster risk reduction ranges from 1:2 to 1:4 (4), but efforts remain underfunded. In particular, inadequate pricing of costs and benefits leads



What if this disaster should happen again? Indian dabbawalas (lunch box carriers) walk through a flooded railway track, after torrential rains paralyzed Mumbai, 27 July 2005. The city's weather bureau said that Mumbai received 944.2 millimeters (37.1 inches) of rainfall in 24 hours.

to inappropriate valuation of investment and financial calculations in risk-reducing measures (5). Risk reduction is not usually referred to as climate adaptation, but may be described as plant breeding and selection, flood-risk reduction, public health care, and so on. Developing countries have many opportunities to integrate climate adaptation in disaster risk-reduction efforts (6). More generally, disaster aid is probably best spent on ex ante risk reduction (7).

Develop and apply innovative finance mechanisms. Industries with greatest exposure have responded to increasing losses with innovative products. Catastrophe bonds are a mechanism used to transfer peak risks to the capital markets, with the range of hazards covered continuing to expand, recently to the flood risks in the U.K. (8). Previously uninsured flood risks in Belgium and the Netherlands are to be covered through public-

Action on disaster risk reduction can support sustainable development under climate change.

private insurance constructions. Existing development financing within local communities, for example, investment funds for small infrastructure improvement in El Salvador, support risk reduction (9), and community groups in India have developed deficit rainfall insurance (9). In Colombia, microentrepreneurs offer affordable and easy to understand life and property microinsurance to the most vulnerable. The World Bank-sponsored Caribbean Catastrophe Risk Insurance Facility offers governments cover against hurricanes and earthquakes with funds available a few days after the event (10). The Munich Climate Insurance Initiative brings together the World Bank, insurers, nongovernmental organizations, and the scientific community to develop finance solutions for adaptation in developing countries (11).

If present trends continue, global disaster losses will keep outpacing average economic growth. Therefore, disaster risk reduction must be core to climate adaptation policies. Numerous mechanisms for action exist that can contribute to the aim of sustainable development.

References and Notes

1. Intergovernmental Panel on Climate Change (IPCC), *Climate Change 2007* (IPCC, Geneva, 2007); www.ipcc.ch.
2. R. A. Pielke Jr., C. N. Landsea, *Bull. Am. Meteorol. Soc.* **80**, 2027 (1999).
3. See Supporting Online Material (SOM).
4. R. Mechler, *Cost-Benefit Analysis of Natural Disaster Risk Management in Developing and Emerging Countries* [GTZ (German Society for Technical Cooperation), Eschborn, 2005].
5. C. Benson, J. Twigg, *Measuring Mitigation: Methodologies for Assessing Natural Hazard Risks and the Net Benefits of Mitigation* (ProVention, Geneva, 2004).
6. R. Few, H. Osbahr, L. M. Bouwer, D. Viner, F. Sperling, *Linking Climate Change Adaptation and Disaster Risk Management for Sustainable Poverty Reduction* [Vulnerability and Adaptation Resource Group (VARG), Washington, DC, 2006]; www.climatevarg.org.
7. J. Linnerooth-Bayer, R. Mechler, G. Pfug, *Science* **309**, 1044 (2005).
8. Allianz, press release, 10 April 2007; www.allianz.com.
9. K. Warner, L. M. Bouwer, W. Ammann, *Environ. Hazards* **7**, 32 (2007).
10. Munich Re, press release, 1 June 2007; www.munichre.com.
11. Munich Climate Insurance Initiative, www.climate-insurance.org.
12. L.M.B. acknowledges support from the Climate Changes Spatial Planning Programme and R.A.P. from the NSF under grant 0345604.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/753/DC1

10.1126/science.1149628

¹Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands. ²Risk Frontiers, Macquarie University, Sydney, Australia. ³Geo Risks Research Department, Munich Re, Munich, Germany. ⁴Center for Science and Technology Policy Research, University of Colorado, Boulder, CO, USA.

*Author for correspondence. E-mail: laurens.bouwer@ivm.falw.vu.nl

Acceptable nAGging

David L. Stocum

The ability to replace lost or damaged body parts varies greatly among animals. Invertebrates, such as sponges, hydra, and flatworms, can regenerate entire organisms from a mere cluster of cells. As for vertebrates, amphibians can regenerate limbs and other complex structures. Despite a rich research history with these model organisms, we are still deciphering how stem cells, cell dedifferentiation (regression to an earlier, unspecialized form), and specific genetic and cellular signaling mechanisms control animal regrowth (see the figure). On page 772 in this issue, Kumar *et al.* (1) report the discovery of a new molecular cue that promotes limb regeneration in newts. Such insights could explain why mammals have limited regrowth abilities, and help guide the field of regenerative medicine.

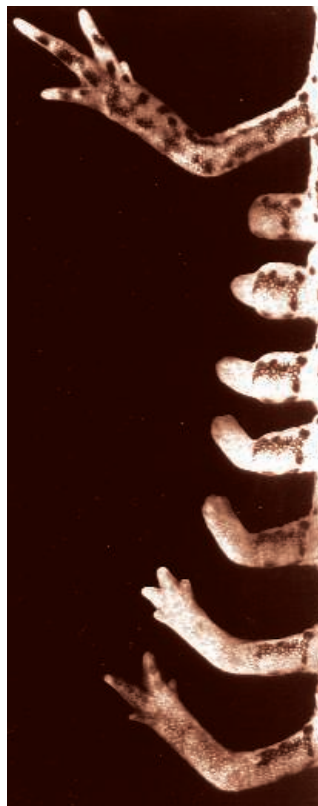
Newt limbs regenerate by dedifferentiating fibroblast, muscle, skeletal, and nonneuronal (Schwann) cells of the nervous system into their earlier stem cell form. These stem cells then proliferate to form a mass of undifferentiated cells, the blastema, in a process that requires factors from both the nerves and overlying wound epidermis (2). Fibroblast growth factor-2 (Fgf-2) (3), glial growth factor-2 (4), the neurotransmitter substance P (5), and the iron transport protein transferrin (6) promote blastema cell proliferation in vitro and in vivo, and Fgf-2-soaked beads support regeneration of denervated axolotl limbs to the point of forming digits when implanted into a blastema that has already attained a critical mass of cells (3). None of these molecules, however, has been put to the rigorous test of supporting regeneration of a denervated and amputated limb from the early dedifferentiation stage when only a few blastema cells are present. Kumar *et al.* pass this test with a previously unknown factor that is expressed by limb nerves.

The new factor is newt anterior gradient (nAG) protein, a homolog of the *Xenopus laevis* secreted protein XAG2, which specifies the development of anterior structures (such as the forebrain) during frog embryogenesis (7). Kumar *et al.* identified nAG by screening

expression libraries from normal limb tissue and blastemas with Prod-1 protein. Prod-1 is a cell surface protein expressed in a proximal-to-distal gradient in the newt limb regeneration blastema that reflects a gradient of cellular positional identity (8). The authors show that nAG is a secreted ligand for Prod-1, and is a mitogen (stimulates proliferation) for newt blastema cells. During the early dedifferentiation stage of newt limb regeneration, nAG was detected in Schwann cells of the distal nerve sheath; later, at the early bud stage, it was expressed in gland cells of the wound epidermis.

But as with all other mitogens for blastema cells, direct in vivo evidence for secretion of nAG from the nerve sheath into the blastema was not demonstrated. However, nAG expression in gland cells required a nerve supply. Most spectacularly, introduction of nAG DNA (by electroporation, a cell permeabilization technique) restored regeneration to digit stages, without promoting nerve regeneration into the wounded limb.

These results raise several questions. Is nAG unique for regenerating limbs, or is its success attributable to its continuous long-term expression, achieved through the gene electroporation strategy? What is the functional relation between the nerve and the wound epidermis? The wound epidermis is also essential for blastema formation, and most likely provides proliferation factors very early after amputation. Because XAG2 signaling in a frog embryo depends on an Fgf signaling pathway (7), might nAG signaling also depend on Fgf-2 from the wound epidermis? Furthermore, as the blastema grows, the distance that nAG diffuses from the distal nerve sheath may be limited. Does additional nAG come from the wound epidermis? And what maintains nAG synthesis by Schwann cells?



Mystery of regeneration. Understanding how vertebrates such as the newt (shown) can regrow body parts may have implications for regenerative medicine. [Adapted from (12)]

Understanding how a factor secreted by neural tissue promotes newt limb regrowth may help the field of regenerative medicine.

The most likely answer to the last question is axons, the cellular extensions of neurons. Schwann cells dedifferentiate and enter the blastema, presumably losing nAG expression, while unmyelinated axons reinnervate the blastema as it grows. As Schwann cells form a myelin sheath around new limb axons, they would be expected to resume axon-induced nAG expression.

A major question is whether nAG-Prod-1 interaction links blastema proliferation with patterning—defining the proximal and distal regions—of new limb parts. Removal of Prod-1 from the blastema cell surface, or its inhibition by Prod-1-specific antibody, abolishes differences in cell adhesion associated with proximodistal positional identity of blastema cells (8), and overexpression of Prod-1 results in the proximal translocation of distal blastema cells transplanted into proximal

blastemas (9). The results of Kumar *et al.* suggest that Prod-1 is likely not directly involved in establishing positional identity, but reflects the process that does. Prod-1 may detect gaps in positional identity, and through its ligand nAG, promotes blastema cell proliferation to fill in such gaps. Organization of these cells is then further defined by other signaling molecules involved in embryonic limb development (10, 11).

The regeneration blastema is a self-organizing structure from its inception, but we don't know how it specifies the pattern of the new limb parts. One idea is that proximal and distal cellular boundaries are set up simultaneously in the early blastema, with intermediate positional identities intercalated between them (2). Solving the mechanism of proximodistal patterning may be the most important key to inducing the regrowth

The author is in the Department of Biology, Indiana University—Purdue University Indianapolis, 723 W. Michigan St. Indianapolis, IN 46202, USA. E-mail dstocum@iupui.edu

of seemingly regeneration-deficient limbs. As Kumar *et al.* state, if we understood how the pattern of the regeneration blastema is specified, we could engineer that pattern into cells of nonregenerating appendages. Growth-promoting molecules could then be supplied exogenously to manifest the pattern without other intervention. How soon this might be possible, particularly in humans, is anyone's guess, but

the addition of nAG to the repertoire of necessary factors is an important step forward.

References

1. A. Kumar, J. W. Godwin, P. B. Gates, A. A. Garza-Garcia, J. P. Brookes, *Science* **318**, 772 (2007).
2. D. L. Stocum, *Regenerative Biology and Medicine* (Academic Press, Elsevier, San Diego and New York, 2006).
3. L. M. Mullen *et al.*, *Development* **122**, 3487 (1996).
4. L. Wang, M. A. Marchionni, R. A. Tassava, *J. Neurobiol.* **43**, 150 (2000).
5. M. Globus, P. Alles, *J. Exp. Zool.* **254**, 165 (1990).
6. A. L. Mescher *et al.*, *Dev. Growth Diff.* **39**, 677 (1997).
7. F. Aberger, G. Weidinger, H. Grunz, K. Richter, *Mech. Dev.* **72**, 115 (1998).
8. S. M. Morais da Silva, P. B. Gates, J. P. Brookes, *Dev. Cell* **3**, 547 (2002).
9. K. Echeverri, E. M. Tanaka, *Dev. Biol.* **279**, 391 (2005).
10. C. Tickle, *Nat. Rev. Mol. Cell Biol.* **7**, 45 (2006).
11. R. E. Hill, *Dev. Growth Diff.* **49**, 439 (2007).
12. J. P. Brookes *et al.* *Science* **276**, 81 (1997).

10.1126/science.1150795

CHEMISTRY

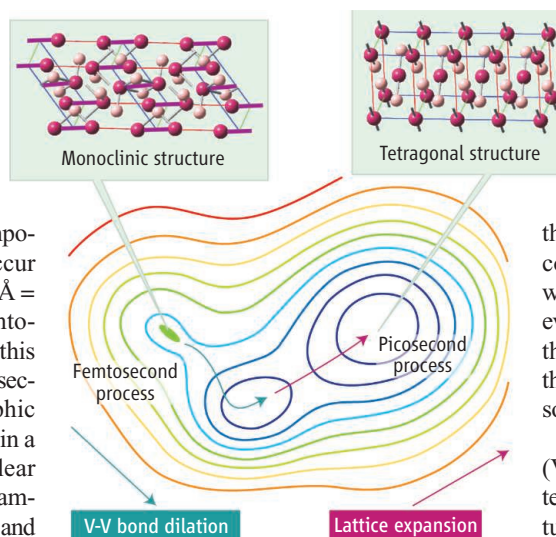
All at Once

Andrea Cavalleri

Complex solids with strongly correlated electrons exhibit transformations in which more than one microscopic property (such as atomic positions, as well as electronic or magnetic arrangements) changes at once. To sort out the underlying physics, one must be working at the spatial and temporal resolution of the change, which can occur over length scales of a few angstroms ($1 \text{ \AA} = 10^{-10} \text{ m}$) and last as little as a few femtoseconds ($1 \text{ fs} = 10^{-15} \text{ s}$). On page 788 of this issue, Baum, Yang, and Zewail use femtosecond electron pulses to take crystallographic snapshots of the evolving lattice structure in a complex solid (1). From their data, a clear connection can be made between the dynamics of a single element of lattice symmetry and a change in electronic properties.

In recent years, complex transition-metal oxides have received widespread attention, in part motivated by the quest to understand high-temperature superconductivity in doped cuprates. Along the way, a wealth of new phenomena have been discovered, including unconventional electronic and magnetic phase transitions, colossal negative magnetoresistance, and novel effects at interfaces. This richness arises from the peculiar physics of partially filled 3d orbitals, which dominate the electronic and magnetic properties of transition metal compounds.

Because of their enormous mutual repulsion, and because they are trapped by atomic distortions, the 3d electrons tend to localize on the transition-metal site, making the system insulating. Yet the same electrons also tend to hybridize with 2p orbitals from their oxygen neighbors and spread over many lattice sites



Coherent transformation. Two-dimensional depiction of the light-induced crystallographic rearrangement in VO_2 studied by Baum and co-workers. At first, the bond between pairs of vanadium atoms dilates rapidly along one axis (left). Only on longer (picosecond) timescales does the system relax along the orthogonal direction, transforming a compressed tetragonal structure to a stable product phase (right).

to minimize their kinetic energy. When this happens, a metal is formed, and magnetic arrangements tend to change as well. Nowhere is this competition more striking than in the abrupt electronic and magnetic transitions that occur when the density of conducting electrons is changed by chemical doping. Similarly, because of the importance of next-neighbor orbital overlap, small distortions in lattice structure can cause large changes in electronic and magnetic properties.

With these considerations in mind, it is easy to see how excitation with light (2)—which rearranges the filling of different orbitals—can transform the structural, electronic, and magnetic properties of a complex

Femtosecond electron crystallography reveals the atomic structural changes that underpin a light-induced insulator-metal transition.

solid (3). Because these rearrangements are cooperative, the whole solid may change when just one photon is absorbed for every 1000 unit cells. However, as we try to understand these bewildering phenomena, a new simplicity emerges on the ultrafast time scale (4). The reason is that the rapid photoexcitation event can impose coherence (5). A coherent change is one where all the unit cells that form the lattice evolve in lockstep with each other, affecting the macroscopic response of the system as they evolve synchronously and allowing us to sort them in the time domain.

The ultrafast physics of vanadium dioxide (VO_2) is a prototypical example. This room-temperature insulator has a monoclinic structure. Pairing and tilting of vanadium pairs along one axis (see the figure, top left) is at the heart of the insulating behavior. Part of the glue that keeps the paired atoms close to one another comes from the localization of two electrons on the vanadium sites, and from the gain in exchange energy when spins point in opposite directions. Photoexcitation removes a fraction of these electrons and slightly weakens a fraction of the bonds—enough to unleash a collective relaxation of the structural distortion, a delocalization of the charge, and a loss of magnetic order. Previous ultrafast optical and x-ray studies showed a rapid, nonequilibrium transition to the metallic phase, accompanied by an equally rapid rearrangement of the atomic structure (6). Yet, in these studies, only a single spot in the diffraction pattern was measured, giving insufficient information to elucidate the changes in atomic structure.

In the work published in this issue, Baum *et al.* were able to detect ultrafast changes in vanadium dioxide in a number of diffraction spots at the same time. The most important information comes in their ability to identify

The author is in the Department of Physics, Oxford University, Oxford OX1 3PU, UK. E-mail: a.cavalleri1@physics.ox.ac.uk

that the fastest atomic motions during this change occur along one particular axis of the crystal, and are related to pairs of vanadium atoms moving apart from one another (see the figure, left). Only at later times do the other crystallographic planes expand (see the figure, right). The transformation pathway between stable monoclinic and tetragonal phases is then shown to pass through an unstable, tetragonal unit cell, which is compressed along one of the axes. The authors thus establish a direct connection between the femtosecond dilation of the V-V bond and the equally fast changes in conductivity that can be measured with other techniques (7).

It is impressive that Baum *et al.* achieve this femtosecond time resolution in the reflection mode of diffraction, because the inherent velocity mismatch between electrons and photons can, in principle, smear out the time response and hinder the observation of the femtosecond movements of the atoms. To this end, the authors ingeniously tilted the front of the optical pump pulse, thus matching the speeds with which the surface of the sample is

excited optically and swept by the diffraction probe. Reflection geometry opens the way to femtosecond electron diffraction in most bulk solids, whereas previous experiments performed in transmission were limited to very thin films (8, 9).

In thinking of new advances in the studies of ultrafast structural dynamics, a few key considerations come to mind. Electron pulses can be incorporated in a microscopy apparatus, as shown previously by Zewail and co-workers (10). This microscopy advance is important for the study of strongly correlated transition-metal oxides discussed here, which are known to exhibit important phase separation phenomena that are quite difficult to investigate (11).

The next frontier will be a time resolution of 10 fs or below, which will allow the movements of the lightest atoms that compose many important organic compounds and liquids to be resolved. Femtosecond electron diffraction is evolving hand in hand with x-ray techniques (12), which have developed in the recent past with both tabletop and accelerator-based techniques, offering

similar time resolutions but also important spectroscopic capabilities (13). The capability of interrogating matter with ultrafast electron and x-ray pulses is opening new horizons that could only be dreamed of as recently as a decade ago.

References

1. P. Baum, D.-S. Yang, A. H. Zewail, *Science* **318**, 788 (2007).
2. S. Koshihara *et al.*, *Phys. Rev. B* **44**, 431 (1991).
3. K. Miyano, T. Tanaka, Y. Tomioka, Y. Tokura, *Phys. Rev. Lett.* **78**, 4257 (1997).
4. N. Gedik, D.-S. Yang, G. Logvenov, I. Bozovic, A. H. Zewail, *Science* **316**, 425 (2007).
5. D. Polli *et al.*, *Nat. Mater.* **6**, 643 (2007).
6. A. Cavalleri *et al.*, *Phys. Rev. Lett.* **87**, 234701 (2001).
7. C. Kübler, *Phys. Rev. Lett.* **99**, 116401 (2007).
8. S. Williamson, G. Mourou, J. C. M. Li, *Phys. Rev. Lett.* **52**, 2364 (1984).
9. B. J. Siwick, J. R. Dwyer, R. E. Jordan, R. J. D. Miller, *Science* **302**, 1382 (2003).
10. V. A. Lobastov, R. Srinivasan, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7069 (2005).
11. E. Dagotto, *Nanoscale Phase Separation and Colossal Magnetoresistance* (Springer, Berlin, 2002).
12. C. Rischel *et al.*, *Nature* **390**, 490 (1997).
13. A. Cavalleri *et al.*, *Phys. Rev. Lett.* **95**, 67405 (2005).

10.1126/science.1150672

CHEMISTRY

No Protection Required

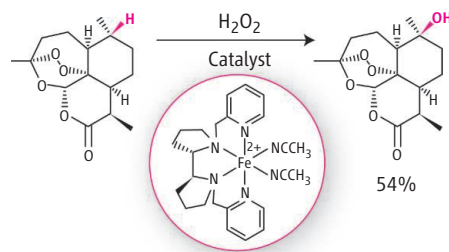
Robert H. Crabtree

Organic synthesis has traditionally relied heavily on activating groups and protecting groups to steer synthetic reactions to the desired products. Activating groups such as halides enhance the reactivity of reactants, whereas protecting groups such as amides or esters block reactivity at undesired sites. On page 783 of this issue, Chen and White (1) show that such activating and protection groups are not always required even in reactions involving complex molecules.

With the rise of green chemistry (2), more attention is being paid to eliminating activating and protecting groups, wherever possible, for two reasons. First, they generate waste. Second, both activating and protecting groups require extra synthesis steps to be introduced into reactants; protecting groups also need extra steps to be removed after reaction. Catalysis can give reactivity and selectivity without the need for activating or protecting groups. The ideal catalyst reacts with an unac-

tivated reactant with a selectivity that can be tuned by choice of catalyst. In practice, however, we are still far from the goal, particularly when the substrate is a complex organic molecule with multiple functional groups.

Perhaps the greatest challenge is finding catalysts that selectively attack C-H bonds,



Selective conversion. The antimalarial compound artemisinin (**left**) is extracted from a shrub used in herbal form in Chinese traditional medicine. Although artemisinin has numerous C-H bonds and a delicate peroxide functional group, it gives a single product (**right**) when the Chen-White catalyst is used in conjunction with hydrogen peroxide. This implies that the catalyst has high selectivity even for a complex molecule, but predictability for other cases will require more detailed study.

An iron catalyst converts C-H bonds to C-OH groups with predictable selectivity even in very large molecules.

which are ubiquitous in organic compounds but are often very unreactive. A number of catalysts are known for this “C-H activation” reaction (3), but they act only on simple molecules such as hydrocarbons. In more complex organic molecules, such as those commonly encountered in pharmaceuticals, numerous oxygen or nitrogen-containing functional groups are distributed over a core held together by carbon-carbon bonds. In such a polyfunctional molecule, unselective attack at any of a number of C-H bonds can result in a cocktail of final products.

Chen and White now report a striking counterexample that shows how C-H bonds can be activated selectively even in complex polyfunctional molecules (see the figure for an example). The authors used an iron catalyst to convert specific C-H bonds in a wide variety of molecules to C-OH groups; the benign and inexpensive hydrogen peroxide serves as the ultimate source of the oxygen atom.

Depending on the specific case, the authors ascribe the remarkably high selectivity to a combination of a number of causes. These include the reactive C-H bond being either inherently more reactive than any other, or

The author is in the Department of Chemistry, Yale University, New Haven, CT 06520, USA. E-mail: robert.crabtree@yale.edu

more physically accessible to the catalyst. The catalyst can also be attracted to a specific location by binding to a pre-existing functional group within the reactant, thus attacking only a nearby C–H bond. A goal in the area is to understand the relevant selectivity trends from the previous results, in order to predict the outcome in any subsequent case. Predictability is essential for the design of a multistep synthetic route relying on a selective, late-stage C–H

activation, because failure at a later step would vitiate the entire scheme.

This remarkable work is part of an emerging trend, in which different types of selective, catalytic C–H activation reactions are being successfully applied to more complex molecules than previously envisaged (4–6). With the conceptual barrier breached for hydroxylation, further striking applications to complex molecules are likely to emerge in the near future.

References

1. M. S. Chen, M. C. White, *Science* **318**, 783 (2007).
2. P. T. Anastas, M. M. Kirchhoff, *Acc. Chem. Res.* **35**, 686 (2002).
3. G. Dyker, Ed., *Handbook of C–H Transformations* (Wiley-VCH, Weinheim, 2005).
4. A. R. Dick, M. S. Sanford, *Tetrahedron* **26**, 2439 (2006).
5. S. Das, C. D. Incarvito, R. H. Crabtree, G. W. Brudvig, *Science* **312**, 1941 (2006).
6. J. A. Labinger, J. E. Bercaw, *Nature* **417**, 507 (2002).

10.1126/science.1150982

SYSTEMS BIOLOGY

A Clock with a Flip Switch

Andy C. Poon and James E. Ferrell Jr.

Two years ago, Takao Kondo's group showed that when a phosphate source (adenosine 5'-triphosphate) and three purified proteins were mixed in a test tube, they spontaneously generated sustained oscillations in the phosphorylation state of one of the proteins (1). The three proteins—KaiA, KaiB, and KaiC—were previously identified as important for the daily patterns of activity and behavior (circadian rhythms) in the cyanobacterium *Synechococcus elongates*. Astonishingly, oscillations of phosphorylation in the reconstituted system were similar to the bacterium's natural circadian rhythm of about 24 hours. Moreover, mutations in one of the proteins, KaiC, that change the circadian period in vivo had nearly identical effects in vitro. Thus, a relatively simple and highly robust timekeeper seemed to set the pace for this complicated organism. The big remaining challenge was to determine how this clock works. Two studies, reported by the Kondo group (2) and by Rust *et al.* on page 809 of this issue (3), now provide a satisfying answer to this question. The oscillations arise from the slow, orderly addition and then subtraction of two phosphates from the KaiC protein. This provides a fascinating example of reductionistic systems biology, where the ability to pick apart a complex system has yielded an understanding of how the whole system works.

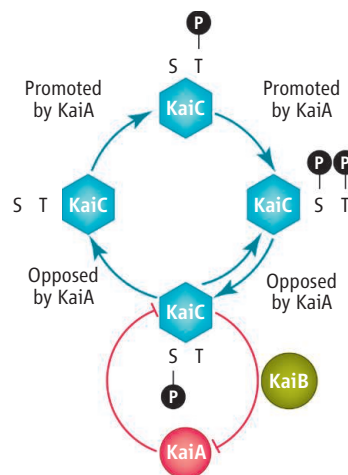
The addition and removal of phosphate can alter a protein's function, and if the protein is part of a network of interacting factors, its phosphorylation status may relay information that impinges on some cell behavior. The dynamics of reversible phosphate addition

and removal in cells are usually rapid—occurring on time scales of seconds or minutes—and so seem poorly suited for slow circadian rhythms. Both studies examine details of the timing of KaiC phosphorylation during oscillations. KaiC is phosphorylated at two sites and in a particular order: first on a threonine residue, then on a serine. Subsequently, the threonine and then serine are dephosphorylated and the KaiC returns to an unphosphorylated state (see the figure). The KaiA protein regulates these transitions by promoting autophosphorylation and inhibiting autodephosphorylation by KaiC.

But a cycle of phosphorylations and dephosphorylations would not necessarily be expected to generate oscillations. Consider, for example, activation and inactivation of the enzyme Erk2 (extracellular signal-regulated protein kinase 2). Like KaiC, Erk2 cycles among four chemical states, modified by enzymes that phosphorylate a tyrosine residue and then a threonine (4), and then by enzymes that dephosphorylate tyrosine first and then the threonine (5). But there is no hint that this system oscillates. What, then, keeps the cyclic phosphorylation and dephosphorylation of KaiC from settling into a static steady-state and allows it to oscillate?

The key insight was the discovery by Rust *et al.* that the serine-phosphorylated form of

The heart of circadian timekeeping in cyanobacteria is a toggle switch that controls the periodic phosphorylation of a key circuit protein.



The cyanobacterial circadian clock. Cyclic phosphorylation and dephosphorylation of the clock protein KaiC on serine (S) and threonine (T) becomes oscillatory through a double-negative-feedback loop (red) that toggles between two states with high or low concentration of free KaiA protein. KaiB is the third protein of the oscillator.

KaiC (S-KaiC) binds stoichiometrically to both KaiA and KaiB. The formation of the KaiA-KaiB-KaiC complex prevents KaiA from activating KaiC phosphorylation. Thus, when S-KaiC concentration is high, KaiA is sequestered by S-KaiC and KaiB, and KaiC dephosphorylation predominates; when S-KaiC concentration is low, KaiA is released and KaiC phosphorylation is activated.

Through modeling studies, Rust *et al.* show that the stoichiometric inhibition of KaiA by S-KaiC allows the cyclic phosphorylation-dephosphorylation system to become an oscillator. This inhibition closes a feedback loop and makes KaiA and S-KaiC mutually antagonistic. KaiA is a negative regulator of S-KaiC, because it pushes the balance between S-KaiC and ST-KaiC (phosphorylated on serine and threonine) toward the latter. Conversely, S-KaiC is a negative regulator of KaiA, because it sequesters KaiA (with the help of KaiB). This mutual antagonism, or double-negative-feedback loop (see the figure), allows S-KaiC and KaiA to function as a bistable toggle switch with two alternative stable steady-states; oscillations could then arise from the successive flipping of the KaiA/S-KaiC switch between these two states. Starting with unphosphorylated KaiC and the KaiA/S-KaiC switch in its low-S-KaiC concentration state, phosphorylation

The authors are in the Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305–5174, USA. E-mail: james.ferrell@stanford.edu

lation will predominate, and more and more ST-KaiC will build up. As the concentration of ST-KaiC increases, S-KaiC will increase too. The increasing S-KaiC will sequester more and more KaiA, which increases the rate of formation of S-KaiC. This then causes more sequestration of KaiA, and so on. Eventually dephosphorylation dominates, and the system is driven back to unphosphorylated KaiC.

What, then, flips the KaiA/S-KaiC switch back to release KaiA and favor phosphorylation? With KaiA sequestered, S-KaiC formation cannot be maintained indefinitely; the rate of S-KaiC production from its immediate precursor ST-KaiC will eventually slow down as ST-KaiC becomes depleted. This effect is the equivalent of a slow negative-feedback loop—an increase in S-KaiC concentration decreases the amount of ST-KaiC, which

decreases the rate of formation of S-KaiC.

Thus, the circadian oscillator system can be thought of as a bistable switch, toggled first by the slow accumulation of S-KaiC and then by a slow negative-feedback loop. This type of circuit can oscillate, as Rust *et al.* demonstrate through a simple differential equation model whose parameters are constrained by their experimental observations.

At first glance, the circadian oscillator of eukaryotes does not seem to work in the same way. It is composed mainly of transcriptional regulators and directed protein degradation (6), rather than a stoichiometrically controlled autophosphorylating adenosine triphosphatase like KaiC, and none of the eukaryotic components have any sequence homology to KaiA, KaiB, or KaiC. On the other hand, the design principles of the two oscillators

may be quite similar. Both circuits include double-negative-feedback loops that might function as bistable triggers, and both include slow negative-feedback loops (8). In terms of systems-level logic, these oscillators appear more similar than different. Maybe this is how a successful circadian oscillator has to be built.

References

1. M. Nakajima *et al.*, *Science* **308**, 414 (2005).
2. T. Nishiwaki *et al.*, *EMBO J.* **26**, 4029 (2007).
3. M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, E. K. O'Shea, *Science* **318**, 809 (2007); published online 4 October 2007 (10.1126/science.1148596).
4. J. E. Ferrell Jr., R. R. Bhatt, *J. Biol. Chem.* **272**, 19008 (1997).
5. Y. Zhao, Z. Y. Zhang, *J. Biol. Chem.* **276**, 32382 (2001).
6. S. M. Reppert, D. R. Weaver, *Nature* **418**, 935 (2002).

10.1126/science.1150740

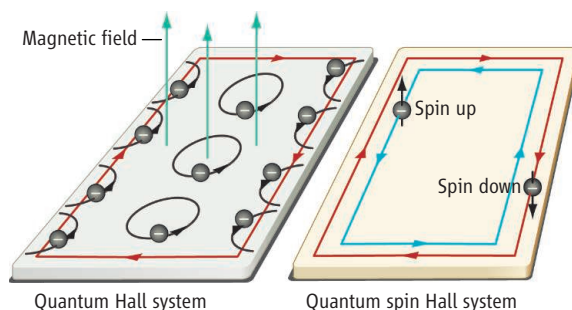
PHYSICS

A New State of Quantum Matter

Naoto Nagaosa

Electrons have a property known as spin, and these spins can be controlled and directed by applied electric and magnetic fields. In recent years, researchers in the relatively young field of spintronics have explored this effect for applications in microelectronics [reviewed in (1)]. The goal is to control and use spins much as today's integrated circuits use the property of electric charge for computing operations. On page 766 of this issue, König *et al.* (2) report experimental results that show the existence of a new state of matter that may take spintronics even further. Not only does this work offer us a look at fundamentally new physical phenomena, it may also allow the development of novel spintronics devices.

Many researchers have explored the so-called spin Hall effect as a possible route to spintronic applications. The original Hall effect goes back to the late 1800s, when Edwin Hall noticed that a voltage would form perpendicularly to a current flowing in a conductor in a magnetic field. In the quantum Hall effects, which were discovered in the 1980 and 1982, the electrical conductance takes on quantized values. In the spin Hall effect, the direction of the flow of electrons can be controlled, depending on whether the spin is up or down, by an



Spin control. In the conventional quantum Hall system (left), the applied magnetic field causes electrons to bounce off the edge of the sample in circular orbits, forming a net flow of charge around the boundary of the material. No magnetic field is needed, however, in the quantum spin Hall system (right), where spin-up and spin-down carriers flow in opposite directions in edge channel states.

applied electric field. The spin Hall effect was proposed theoretically long ago based on an “extrinsic” mechanism (3), in which impurities in a material deflect the spin-up and spin-down electrons in opposite ways. However, recent interest has centered on an intrinsic form of spin Hall effect and the possibility of spin flow without energy dissipation.

Instead of impurities, the intrinsic mechanism relies on the interplay of the spin and orbital motion of the electrons in the perfect periodic background of the crystal lattice, leading to different paths for the up and down spins. The theoretical proposals for this effect (4, 5) were followed by the experimental discovery of the spin Hall effect in GaAs (6, 7)

Experiments show that electron spins can flow without dissipation in a novel electrical insulator.

and also in metallic systems (8–10). However, the detailed mechanism in these cases still needs to be scrutinized; the flow of charges can still be distorted by impurity scattering and thus contribute to the spin current. Therefore, we need to obtain clear observations of a truly dissipationless spin current.

Dissipationless flow is known to occur in the conventional quantum Hall effect, where the electrons are deflected into circular paths by a magnetic field (see the left panel of the figure).

This motion is not random among electrons but is coherently organized, leading to a collective state. This state is stable because a finite amount of energy (the energy gap) is required to disturb it. In a real sample, the electrons bounce back from the edges, causing a net one-dimensional motion, which corresponds to a flow of current around the edge. Because the direction of this motion is one-way, it cannot be scattered backward, and so the dissipation of the flow of charge is suppressed in the quantum Hall system.

A crucial question is whether a similar state is possible for the spin current. My colleagues and I studied this question theoretically in 2004 (11), and we considered the spin

The author is in the Department of Applied Physics, University of Tokyo, Tokyo 113-8656, Japan. E-mail: nagaosa@appi.t.u-tokyo.ac.jp

Hall conductivity for a band insulator with a finite band gap caused by the spin-orbit interaction. In this case, the electrons have a velocity transverse to the external electric field and a direction that depends on the spin. That is, the spin currents in the totally occupied bands do not cancel each other out, and finite spin Hall conductivity results even in the band insulator. This state was called a spin Hall insulator, and the candidate materials we proposed included HgTe, HgSe, HgS, PbTe, PbSe, and PbS. However, it was not clear how to fundamentally distinguish these materials from the usual band insulators.

Kane and Mele (12, 13) achieved a breakthrough when they invented a model for graphene with spin-orbit interaction, which revealed an insulating state with robust helical edge modes, i.e., the modes with opposite spins have opposite directions of propagation. When the number of these helical edge modes is even, some perturbation induces the hybridization of these helical modes and hinders their propagation. This does not occur for an odd number of modes, where the propagation is stable and protected (see the right panel of the figure). The latter case corresponds to a new class of band insulator, i.e., the quantum spin Hall system, and the former to the usual insulator (12, 13).

Soon after these results, Bernevig *et al.* (14) proposed a different way to create a quantum spin Hall system. They looked at the problem as a phase transition between the conventional insulator and the quantum spin Hall system. It turned out that the phase transition is accompanied by a “band crossing,” which changes the number of helical edge modes. Furthermore, they proposed a specific system, i.e., the quantum well of CdTe/HgTe/CdTe, where this phase transition could be induced by changing the thickness of the HgTe layer. This was a strong proposal and appealing enough to motivate experimentalists to try testing it.

König *et al.* now report the experimental observation of these robust helical edge modes. They fabricated quantum well structures and changed the thickness d of the HgTe layer. As d increases, the energy gap decreases and eventually collapses at a critical thickness of $d_c = 6.3$ nm. This gap closing corresponds to the quantum phase transition between the usual insulator at $d < d_c$ and quantum spin Hall state at $d > d_c$. König *et al.* confirmed this by measuring the expected quantized charge conductance consistently with a helical mode for each of the two edges, while observing a much smaller conductance at smaller thickness, indicating the usual insulating state.

This conclusion is further reinforced by the magnetic field dependence of the charge con-

ductance. The conductance should show a rapid decrease as the magnetic field B increases, which König *et al.* also observed experimentally. In this case, the authors did not directly measure the spin Hall conductance, which is not “quantized.” This is because, unlike the charge, the total spin is not a conserved quantity in the presence of the spin-orbit interaction. Therefore, although the search for the intrinsic spin Hall effect in the insulator leads to a new classification of the electron states in solids, the implications for the magneto-transport properties still remain to be studied.

The impact of this work will be far-reaching because it has revealed that there are fundamentally different kinds of insulators. Even after 80 years, the band theory of materials still has new and surprising aspects. A more complete classification scheme for these unusual insulating states is now being constructed, including the three-dimensional systems (15). Eventually, the quantum spin Hall system might enable the design of spin current circuits without dissipation, which will open up new possibilities in spintronics.

References and Notes

1. S. A. Wolf *et al.*, *Science* **294**, 1488 (2001).
2. M. König *et al.*, *Science* **318**, 766 (2007); published online 20 September 2007 (10.1126/science.1148047).
3. M. I. D'yakonov, V. I. Perel', *JETP Lett.* **13**, 467 (1971).
4. S. Murakami, N. Nagaosa, S.-C. Zhang, *Science* **301**, 1348 (2003).
5. J. Sinova *et al.*, *Phys. Rev. Lett.* **92**, 126603 (2004).
6. Y. K. Kato, R. C. Myers, A. C. Gossard, D. D. Awschalom, *Science* **306**, 1910 (2004).
7. J. Wunderlich, B. Kaestner, J. Sinova, T. Jungwirth, *Phys. Rev. Lett.* **94**, 047204 (2005).
8. S. O. Valenzuela, M. Tinkham, *Nature* **442**, 176 (2006).
9. E. Saitoh, M. Ueda, H. Miyajima, G. Tatara, *Appl. Phys. Lett.* **88**, 182509 (2006).
10. T. Kimura *et al.*, *Phys. Rev. Lett.* **98**, 249901 (2007).
11. S. Murakami, N. Nagaosa, S.-C. Zhang, *Phys. Rev. Lett.* **93**, 156804 (2004).
12. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
13. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).
14. B. A. Bernevig, T. L. Hughes, S.-C. Zhang, *Science* **314**, 1757 (2006).
15. L. Fu, C. L. Kane, *Phys. Rev. B* **76**, 045302 (2007).
16. I acknowledge financial support from the Grant-in-Aids under grant numbers 15104006, 16076205, and 17105002, and National Research Grid Initiative Nanoscience Project from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

10.1126/science.1150199

ASTRONOMY

Mining for the Ephemeral

Geoffrey C. Bower

A new generation of telescopes is helping researchers explore transient energetic processes outside our Galaxy.

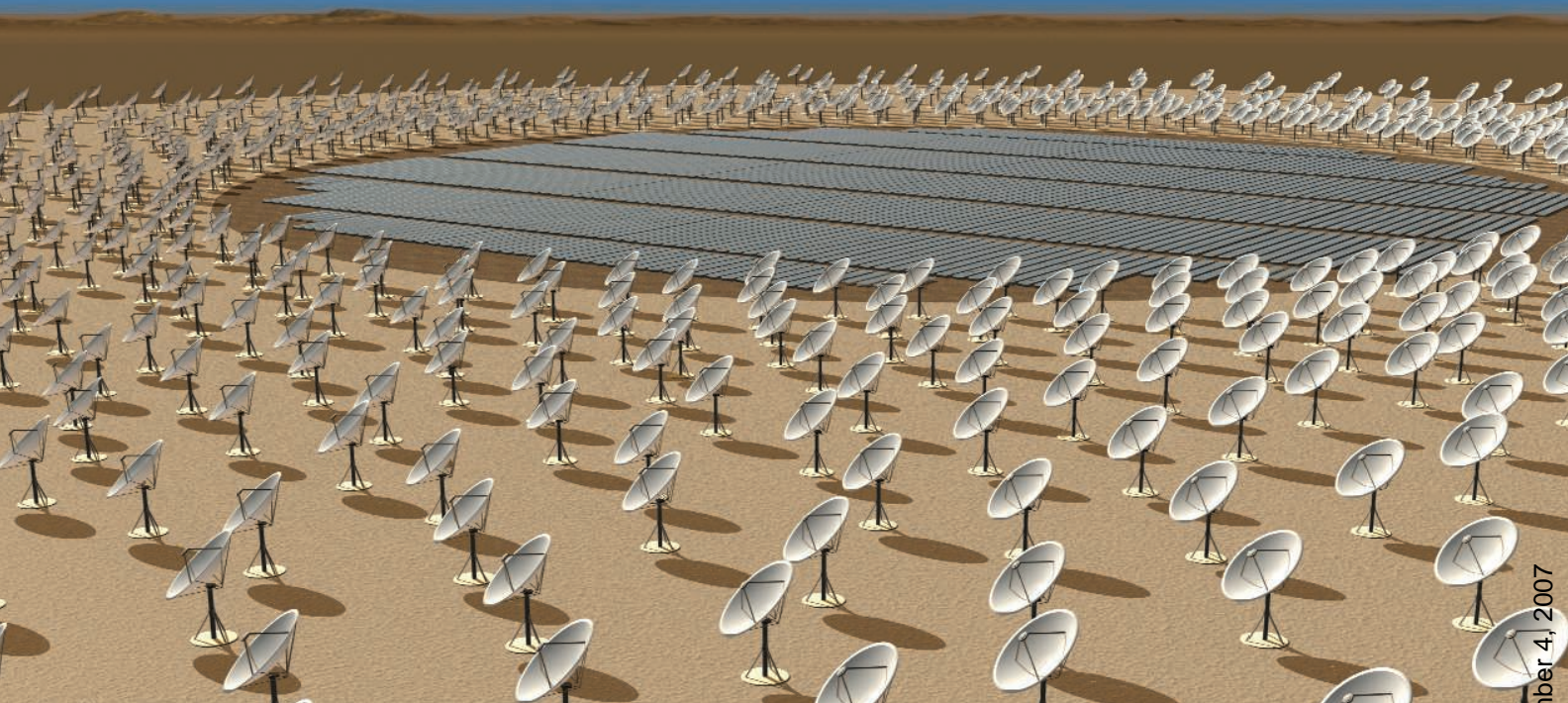
Observed at radio wavelengths, the sky has revealed in recent years that it plays host to a zoo of variable and transient sources that pulse, flicker, burst, and burp. In the most recent manifestation of such ephemeral phenomena, Lorimer *et al.* (1) report on page 777 of this issue the discovery of a single radio burst so intense that it overloaded the detector of the Parkes Radio Telescope in Australia.

The authors argue that the burst originated outside of the Galaxy, possibly at a distance of more than 1 billion light years. This indicates an enormous and unprecedented luminosity as well as possibly providing a new method for studying the intergalactic medium (IGM), one of the most poorly characterized constituents of the universe. More detections could give the first complete census of baryons (i.e., particles such as protons and

neutrons) in the IGM, which is believed to account for 90% of the total baryons in the universe. On the basis of the small sliver of sky and the limited sampling time that led to this discovery, it appears that hundreds of similar events occur throughout the sky every day. The burst heralds a new era of transient discovery at radio and other wavelengths, driven by new telescopes and advanced technology for performing and analyzing enormous surveys of the sky.

In the past century, astronomers escaped the limited spectrum visible to the human eye and developed instruments capable of observing wavelengths from radio waves to gamma rays. As a result, researchers have made numerous discoveries including the microwave background, pulsars, massive black holes, and gamma-ray bursts. Although the static sky has not been fully explored, the objects that vary over time represent the new terra incognita of astronomy. The available parameter space is vast: Transient events are observed at all wavelengths, in a wide variety

The author is in the Department of Astronomy, University of California, 601 Campbell Hall, Berkeley, CA 94720, USA. E-mail: gbower@astro.berkeley.edu



Scanning the skies. The Square Kilometre Array will begin operations in 2020. The facility will have flat arrays or dipole antennas for the low-frequency bands and steerable dishes for medium- and-high frequency bands.

of cosmic environments from star-forming regions to distant galaxies, and on time scales as short as nanoseconds. Thus far, systematic searches have only probed narrow slices of this multidimensional transient parameter space. This is especially true at radio wavelengths because of the small number of large telescopes and the high cost of telescope resources necessary to conduct extensive transient surveys (2).

This newly discovered extragalactic burst illustrates some of the challenges and opportunities for transient science. The burst resembles a phenomenon known as rotating radio transients (RRATs) (3). The discovery of RRATs also required sifting through petabytes of data for a handful of elusive pulses. Both discoveries were made possible by the addition of a multibeam receiver to the Parkes telescope that improved survey speed by an order of magnitude, and by access to high-speed computing. The sporadic RRAT bursts are hypothesized to originate from unusual pulsars (rotating neutron stars that produce periodic emission) located in the Milky Way.

Both RRATs and the extragalactic burst were discovered through a reanalysis of data obtained for a large-scale pulsar survey. Given the similarities in their properties, the extragalactic burst could have been found in the data mining that discovered RRATs. The initial searches for bursts, however, were constrained to cover only a narrow range of a parameter known as dispersion measure, which characterizes the effect that interstellar and intergalactic plasma has on a propagating radio wave. Extending the search

space to include extragalactic targets was computationally expensive and was therefore excluded in order to optimize the search for galactic objects.

In recent years, astronomers have found other classes of radio transient events through analysis of archival data (4) as well as through surveys dedicated to transient discovery (5). These results point to the richness of such phenomena and the richness of existing archives from major facilities such as the Parkes telescope and the Very Large Array.

When a transient is uncovered in the archives, identification of the progenitor object or host can be impossible. Typically, contemporaneous data at multiple wavelengths are essential for determining the underlying physical properties of the transient, as was the case for gamma-ray burst afterglows (6). Further, a radio telescope's capacity to survey large areas of the sky is most easily coupled to low angular resolution; that is, large fields of view are traded off for poor localization of sources. In the case of this burst, the burst is localized to a region one-quarter the size of the Moon, making unique identification very unlikely.

In the coming years, new radio telescopes will dramatically extend our capabilities for transient searches. All are interferometric arrays that simultaneously provide wide fields of view and accurate source localization. The Allen Telescope Array, which recently began operation, makes use of a large array of small dishes for unprecedented survey capabilities at centimeter wavelengths (7). An upgrade to the Very Large Array makes this an effectively

new telescope (the Expanded Very Large Array) with an order-of-magnitude increase in survey speed over a broad wavelength range. Additional projects with improved survey capability are under way in the southwestern United States, the Netherlands, Australia, and South Africa. As powerful as they are, these facilities will be dwarfed by the Square Kilometre Array (see the figure), the next-generation radio telescope with 100 times the sensitivity and a broad array of powerful signal-processing capabilities (8).

The ultimate prize that these telescopes and technologies offer is the discovery of astrophysical phenomena. Pulsars, supernovae, and gamma-ray bursts all represent classes of transient sources that have been important probes of a broad variety of physical processes, including cosmology, galaxy evolution, the basic properties of dense star matter, and general relativity. The burst found by Lorimer *et al.* is a reminder that startling discoveries are still to be made.

References

1. D. R. Lorimer, M. Bailes, M. A. McLaughlin, D. J. Narkevic, F. Crawford, *Science* **318**, 777 (2007); published online 27 September 2007 (10.1126/science.1147532).
2. J. M. Cordes, T. J. W. Lazio, M. McLaughlin, *New Astron. Rev.* **48**, 1459 (2004).
3. M. McLaughlin *et al.*, *Nature* **439**, 817 (2006).
4. G. C. Bower *et al.*, *Astrophys. J.* **666**, 346 (2007).
5. S. Hyman *et al.*, *Nature* **434**, 50 (2005).
6. J. van Paradijs, S. Kouveliotou, R. A. M. J. Wijers, *Annu. Rev. Astron. Astrophys.* **38**, 379 (2000).
7. D. DeBoer *et al.*, *Proc. SPIE* **5489**, 1021 (2004).
8. C. Carilli, S. Rawlings, *New Astron. Rev.* **48**, 979 (2004).

10.1126/science.1149084

The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race

Alexei A. Aravin,* Gregory J. Hannon,† Julius Brennecke*

Increasingly complex networks of small RNAs act through RNA-interference (RNAi) pathways to regulate gene expression, to mediate antiviral responses, to organize chromosomal domains, and to restrain the spread of selfish genetic elements. Historically, RNAi has been defined as a response to double-stranded RNA. However, some small RNA species may not arise from double-stranded RNA precursors. Yet, like microRNAs and small interfering RNAs, such species guide Argonaute proteins to silencing targets through complementary base-pairing. Silencing can be achieved by corecruitment of accessory factors or through the activity of Argonaute itself, which often has endonucleolytic activity. As a specific and adaptive regulatory system, RNAi is used throughout eukarya, which indicates a long evolutionary history. A likely function of RNAi throughout that history is to protect the genome from both pathogenic and parasitic invaders.

Argonaute proteins, in complex with distinct classes of small RNAs, form the core of the RNA-induced silencing complex (RISC), the RNA-interference (RNAi) effector complex (1). The Argonaute superfamily segregates into two clades, the Ago clade and the Piwi clade (table S1). The single fission yeast Argonaute and all plant family members belong to the Ago clade, whereas ciliates and slime molds contain members of the Piwi clade. Together, these findings indicate that Piwis and Agos are similarly ancient. Animal genomes typically contain members of both clades, and it is becoming clear that this division of Argonautes reflects their underlying biology.

Ago clade proteins complex with microRNAs (miRNAs) and small interfering RNAs (siRNAs), which derive from double-stranded RNA (dsRNA) precursors (1). miRNA-Ago complexes reduce the translation and stability of protein-coding mRNAs, which results in a regulatory network that impacts ~30% of all genes. siRNAs in *Drosophila* arise from replicating RNA viruses and are crucial in antiviral immune responses (2). In *Caenorhabditis elegans*, endogenous siRNAs overlap protein-coding genes and likely participate in gene regulation (3). As classes, neither virus-derived nor endogenous siRNAs have yet been described in vertebrates.

The Piwi Clade

The Piwi clade is found in all animals examined so far, and its presence is tightly correlated with the emergence of specialized germ cells. Most

animals separate germline and somatic cells early in development and restrict Piwi expression specifically to germ cells. In flatworms, an animal clade close to the bilaterian root, Piwis are expressed in germ cells and neoblasts, undifferentiated stem cells responsible for the remarkable regenerative capacity of these organisms (4). Although neoblasts are considered somatic stem cells, they are capable of giving rise to germ cells. Thus, the conserved expression pattern of Piwi proteins is a strong indication of their vital function in the germ line.

The genomes of multicellular animals encode multiple Piwi proteins. The three *Drosophila* proteins Piwi, Aubergine, and AGO3 are expressed in the male and female germ lines. Piwi is additionally expressed in the somatic cells, which are in close contact with germline cells (5–8). Expression of the three mouse proteins MIWI (PIWIL1), MILI (PIWIL2), and MIWI2 (PIWIL4) is mainly restricted to the male germ line (9–12). Although expression of *Mili* in prenatal ovaries has been reported (9), no function for Piwis in the female mammalian germ line has yet been demonstrated.

Consistent with their expression pattern, Piwi mutant animals exhibit defects in germ cell development. *Drosophila* Piwi is required for the maintenance of germline stem cells, both in testes and ovaries (13). In mouse, all three Piwi proteins are nonredundantly required for spermatogenesis (10–12). Although some somatic expression of Piwis has been reported, mutant animals lack obvious defects in the soma. On the basis of their loss-of-function phenotypes, Piwi proteins were placed in signaling pathways underlying germline development (10, 14). However, genetic studies also pointed to a role for the Piwi pathway in silencing selfish genetic elements (15–17). Insight into the molecular

function of Piwi proteins was stalled until the discovery of their small RNA partners.

Piwi-Interacting RNAs

The first indication of a distinct population of Piwi-associated small RNAs came from studies in *Drosophila*. The presence of 25- to 27-nucleotide (nt) RNAs homologous to the repetitive *Stellate* locus was correlated with its silencing and required the Piwi clade protein Aubergine (15). Profiling of small RNAs through *Drosophila* development placed *Stellate*-specific small RNAs into a broader class, derived from various repetitive elements, called repeat-associated small interfering RNAs (rasiRNAs) (18). A direct interaction between rasiRNAs and Piwi proteins was demonstrated by immunoprecipitation of Piwi complexes (5–7, 16). Small RNAs resembling *Drosophila* rasiRNAs have been identified in testes and ovaries of zebrafish, which demonstrates evolutionary conservation of this small RNA class (19). Small RNA partners of Piwi proteins were also identified in mammalian testes and termed Piwi-interacting RNAs (piRNAs) (20). Although these RNAs share some features with rasiRNAs, there are also substantial differences, including a dearth of sequences matching repetitive elements. Nonetheless, on the basis of their common features, we refer to small RNAs in Piwi complexes as piRNAs with rasiRNAs being one specialized subclass.

Piwis and piRNAs form a system distinct from the canonical RNAi and miRNA pathways. No association between Piwis and miRNAs was detected in either fly (5, 6) or mouse (21, 22), although piRNAs, like miRNAs, carry a 5' monophosphate group and exhibit a preference for a 5' uridine residue (21–23). In contrast to miRNAs, many of which are conserved through millions of years of evolution, individual piRNAs are poorly conserved even between closely related species (21–23). piRNAs in *Drosophila* (5, 6) and mammals (21–23), as well as siRNA-like scan RNAs that bind Piwi proteins in ciliates (24), are substantially longer (24 to 30 nt) than miRNAs and siRNAs (21 to 23 nt). Unlike animal miRNAs, but similar to plant miRNAs, piRNAs carry a 2'-O-methyl modification at their 3' ends, which is added by a Hen-1 family RNA methyltransferase (25). Finally, genetic analyses in flies (16) and zebrafish (19) argue against a role for Dicer, a key enzyme in miRNA and siRNA biogenesis, in piRNA production.

The Genomic Origin of piRNAs

Most *Drosophila* piRNAs match repetitive elements and therefore map to the genome in dozens to thousands of locations. Yet mapping of those piRNAs that could be placed uniquely in the genome (e.g., piRNAs from divergent repeat copies) identified a limited set of discrete loci that could give rise to most piRNAs. These were dubbed piRNA clusters (5). piRNA clusters range from several to hundreds of kilobases

Watson School of Biological Sciences, Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA.

*These authors contributed equally to the work.

†To whom correspondence should be addressed. E-mail: hannon@cshl.edu

in length. They are devoid of protein coding genes and instead are highly enriched in transposons and other repeats (fig. S1). The vast majority of transposon content in piRNA clusters occurs in the form of nested, truncated, or damaged copies that are likely not capable of autonomous expression or mobilization. The presence of transposable elements per se is not sufficient for piRNA production. Virtually all piRNA clus-

repeats, and even those that do match annotated transposons are diverged from consensus, potentially active copies (fig. S1). Prepachytene piRNAs are found in germ cells before meiosis (26). These share the molecular characteristics of pachytene piRNAs but originate from a different set of clusters that more closely match those of *Drosophila* and zebrafish in repeat content.

complexes, and that is the only family member that is present in the somatic cells of the ovary (5), where *flamenco* is predominantly expressed.

Unlike trans-acting siRNAs in plants, piRNAs do not arise from clusters in a strictly phased manner but rather originate from irregular positions forming pronounced peaks and gaps of piRNA density (Fig. 1). piRNA populations are extremely complex, with our recent estimates placing the number of distinct mammalian pachytene piRNAs at >500,000.

Biogenesis of piRNAs

The lack of a dependence on Dicer (16, 19) and the profound strand asymmetry of mammalian pachytene clusters indicate that piRNAs are not generated from dsRNA precursors. In *Drosophila*, most piRNA clusters generate small RNAs from both strands; however, there are exceptions, such as the *flamenco* locus, where piRNAs map almost exclusively to one genomic strand (fig. S1) (5). In zebrafish, piRNAs can map to both genomic strands; however, within any given region of a cluster, only one strand gives rise to piRNAs (19).

Given these considerations, two plausible models emerge. The first is the generation of piRNAs by sampling of long single-stranded precursors. Alternatively, piRNAs could be made as primary transcription products. Evidence for the former is the lack of a 5' triphosphate group and the observation that a single P-element insertion at the 5' end of the *flamenco* cluster prevents the production of piRNAs up to 160 kb away (5). This strongly supports a model in which a single transcript traverses an entire piRNA cluster and is subsequently processed into mature piRNAs.

Processing of small RNAs from long single-stranded transcripts is not unprecedented. Indeed, miRNAs are processed from precursors that often span several kilobases and that can encode several individual miRNAs (27). Pronounced peaks in piRNA density within a cluster also hint at the existence of specific processing determinants; however, the nature of these signals is yet to be resolved.

The machinery that produces piRNAs from cluster-derived transcripts must be somewhat flexible, as different Piwi proteins in flies and mammals each incorporate a distinct size class of small RNA (5, 21, 22). Data from flies and mammals suggest a model in which piRNA production begins with single cleavage of a primary piRNA cluster transcript to generate a piRNA 5' end. piRNAs may be sampled virtually from any position within a cluster with the only preference being a 5' uridine residue. After incorporation of the cleaved RNA into a Piwi, a second activity generates the 3' end of the piRNA with the specific size determined by the footprint of the particular family member on the RNA.

Piwi and Aubergine complexes contain piRNAs antisense to a wide variety of *Drosophila*

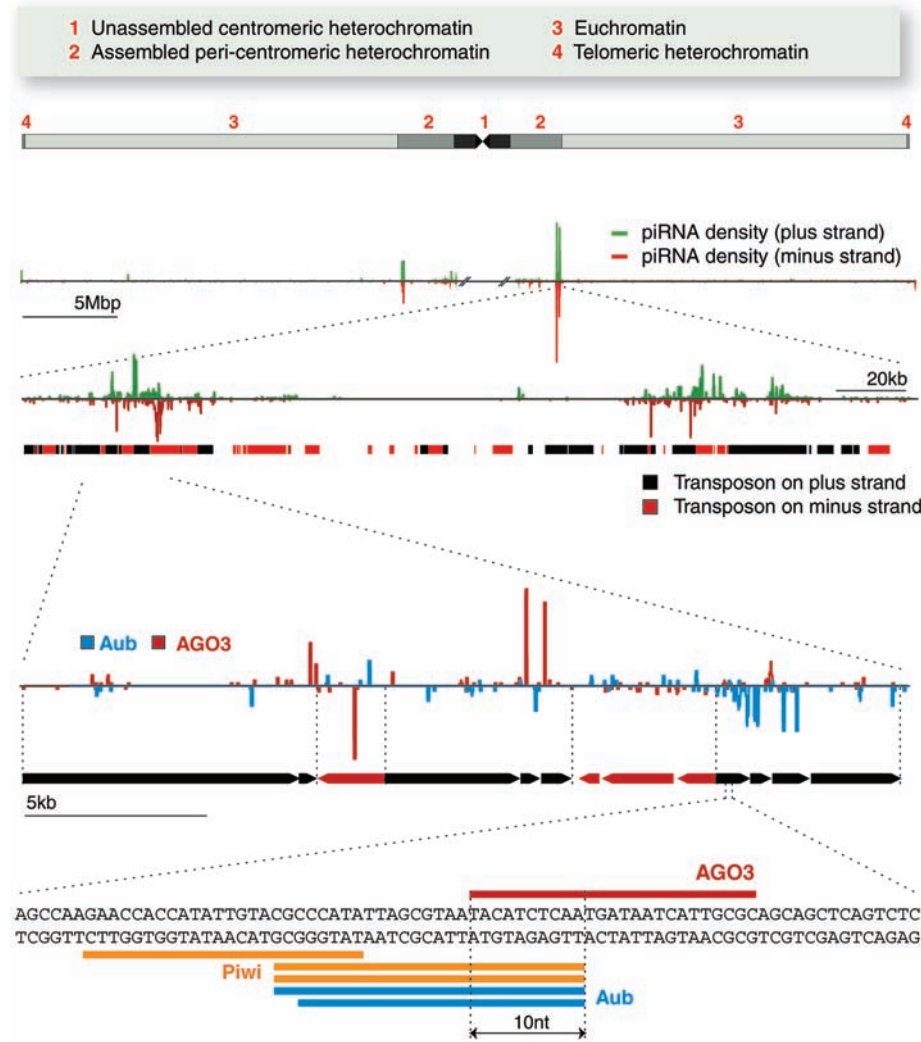


Fig. 1. Features of piRNA clusters. Overview of *Drosophila* chromosome 2 and progressively more detailed view of a major piRNA cluster.

ters in *Drosophila* are located in pericentromeric or telomeric heterochromatin, which suggests that chromatin structure may play a role in defining piRNA clusters (Fig. 1).

Prominent piRNA loci are also found in mammals (21–23, 26) and zebrafish (19). Mammalian piRNAs can be divided into two populations. Pachytene piRNAs appear around the pachytene stage of meiosis, become exceptionally abundant, and persist until the haploid round spermatid stage, after which they gradually disappear during sperm differentiation (21, 22). Pachytene piRNAs are relatively depleted of

Generally, clusters in flies and vertebrates give rise to piRNAs that associate with multiple Piwi proteins. Mouse pachytene piRNAs join both MILI and MIWI complexes (27). Similarly, *Drosophila* clusters produce piRNAs, which associate with all three Piwi proteins (5). However, some clusters generate piRNAs that join specific Piwi proteins, likely because these clusters and the Piwi proteins with which their products associate display specific temporal and special expression patterns. For example, *Drosophila* piRNAs originating from the *flamenco* cluster are found almost exclusively in Piwi

ila transposons, and these show the strong 5' U preference noted for mammalian piRNAs (5, 6, 16). In contrast, AGO3 associates with piRNAs strongly biased toward the sense strand of transposons and with no 5' nucleotide preference (5, 7). piRNAs in AGO3 show a characteristic relation with piRNAs found in Aub complexes, with these small RNAs overlapping by precisely 10 nt at their 5' ends (Fig. 2A). Accordingly, the AGO3-bound piRNAs were strongly enriched for adenine at position 10, which is complementary to the 5' U of Aub-bound piRNAs (5, 7). These observations indicated the existence of two distinct piRNA populations, possibly with different biogenesis mechanisms, and led to the hypothesis that cluster-derived transcripts and transcripts from active transposons interact through the action of Piwi proteins to form a cycle that amplifies piRNAs that target active mobile elements (5).

The cycle (called the Ping-Pong amplification loop) (Fig. 2B) begins with a transposon-rich piRNA cluster giving rise to a variety of piRNAs. In most clusters, a random arrangement of transposon fragments would initially produce a mixture of sense and antisense piRNAs, likely populating Piwi and Aub. When encountering a complementary target, a transposon mRNA, Piwi/Aub complexes cleave 10 nt from the 5' end of their associated piRNA (6, 7). This not only inactivates the target but also creates the 5' end of new AGO3-associated piRNA. Loaded AGO3 complexes are also capable of cleaving complementary targets (7); one place from which such targets could be derived is the clusters themselves. Cleavage of cluster transcripts by AGO3 would then generate additional copies of the original antisense piRNA, which would enter Aub and become available to silence active transposons. The combination of these steps can form a self-amplifying loop. Signatures of this amplification loop are also apparent in zebrafish (19) and in mammalian prepachytene piRNAs (26). This transposon-silencing pathway, with both genetically encoded and adaptive components, has many conceptual similarities to adaptive immune responses.

Function of Piwi Proteins and piRNAs

Studies of piRNAs have pointed to a conserved function in the control of mobile genetic elements, and this is consistent with the defects in transposon suppression observed in Piwi mutants (15–17). One line of evidence comes from studies of loci that exert control over specific transposons in flies. The *flamenco* locus represses transposition of the retrotransposons *gypsy*, *ZAM*, and *Idefix* (28, 29). *flamenco* maps to the pericentromeric heterochromatin on the X chromosome. Genetic analysis failed to reveal a protein-coding gene underlying *flamenco* function; however, the discovery that *flamenco* is a major piRNA cluster provided a molecular basis for its ability to suppress several unrelated retroelements. *flamenco* spans at least 180 kb

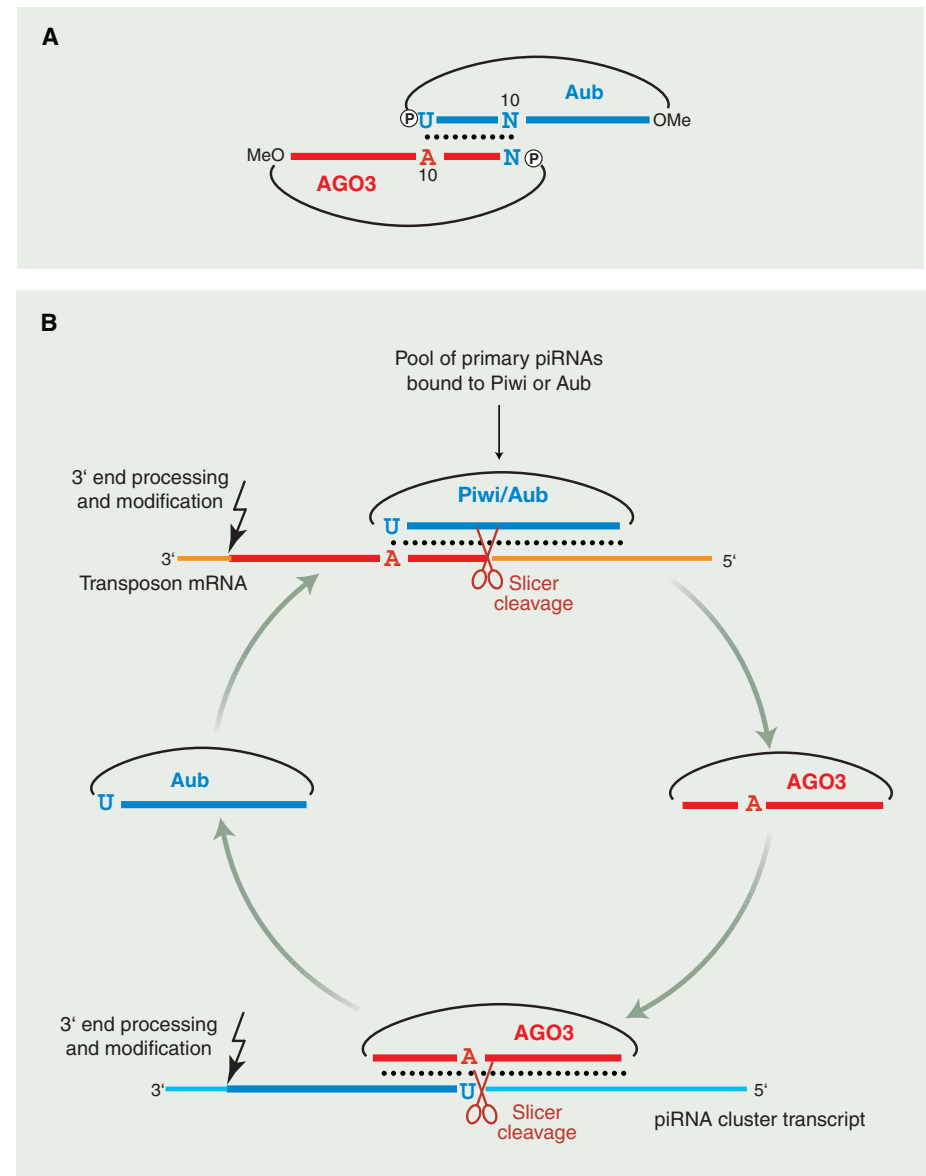


Fig. 2. Properties and biogenesis of piRNAs. **(A)** Features of Aub- and AGO3-associated piRNAs in *Drosophila*. Indicated are the 5' U bias in Aub-bound piRNAs, the 10A bias in AGO3-bound piRNAs, the 5' phosphate, and the 3' O-methylation. **(B)** Ping-Pong model of piRNA biogenesis in *Drosophila*. Primary piRNAs are generated by an unknown mechanism and/or are maternally deposited. Those with a target are specifically amplified via a Slicer-dependent loop involving AGO3 and Aub.

and is highly enriched in many types of repetitive elements, including multiple fragments of *gypsy*, *ZAM*, and *Idefix*. In *flamenco* mutants, *gypsy* is desilenced, and essentially all piRNAs derived from this cluster are lost (5). Thus, *flamenco* is an archetypal piRNA cluster that encodes a specific silencing program, which is parsed by processing into individual, active small RNAs that exert their effects on loci located elsewhere in the genome.

Genetic studies of Piwi mutants suggested involvement in germline development in both invertebrates and vertebrates (8, 10–12, 30). *Drosophila piwi* is required in germ cells, as well as in somatic niche cells, for regulation of cell division and maintenance of germline stem

cells (8). The *aubergine* phenotype resembles so-called spindle-class mutants that demonstrate meiotic progression defects (30). The defects in spindle-class mutants are a direct consequence of Chk2 and ATR (ataxia telangiectasia mutated and Rad3-related) kinase dependent meiotic checkpoint activation, and the phenotypes of *aub* mutants are partially suppressed in animals defective for this surveillance pathway (31).

In mice, loss of individual Piwi proteins causes spermatogenic arrest (10–12). In *Miw1* mutants, germ cells are eliminated by apoptosis after the haploid, round spermatid stage (10). However, in *Mili* (11) and *Miw2* (12) mutants, earlier defects appear as meiosis is arrested

around the pachytene stage. In flies, mammals, and zebrafish, no phenotypic abnormalities have yet been detected outside of the germ line, in accord with the expression pattern of Piwis.

A key question is whether the diverse effects of Piwi mutations can be explained solely through the actions of Piwi proteins in transposon control or whether other Piwi functions exist. In *Drosophila*, studies of hybrid dysgenesis linked transposon activation to severely impaired gametogenesis. Mutation of a single piRNA cluster, *flamenco*, results in defects in germ and follicle cell development and complete sterility (32). Defects in *aub* mutants are linked to DNA damage checkpoint signaling that is probably activated in response to double-strand breaks arising from transposon activity (31). In mammals, germ cell loss in *Mili* and *Miwi2* mutants has been correlated with transposon activation (12, 26). Other studies also support the idea that severe defects in germ cell development can be a direct consequence of transposon activation. For example, *Dnmt3L*-deficient animals show demethylation of transposable elements, which lead to their increased expression, as well as meiotic catastrophe and germ cell loss (33), a combination of phenotypes similar to those seen in *Mili* and *Miwi2* mutants.

Overall, genetic and biochemical data indicate that a substantial component of Piwi biology is dedicated to transposon control. However, there are also properties of the Piwi pathway that are difficult to explain solely on the basis of transposon regulation. Pachytene piRNAs in mammals are depleted of transposon sequences, and even those that form part of this population are highly diverged and unlikely to function in transposon suppression. Consistently, no activation of transposons has been detected in *Miwi* mutants (27). Thus, the function of pachytene piRNAs remains a mystery, as does the basis of postmeiotic arrest in the *Miwi* mutants. Global translational control plays important roles during mammalian spermatogenesis, with the expression of many mRNAs being posttranscriptionally regulated. Loss of *Miwi* has been linked to changes in the abundance of several mRNAs important for development of haploid cells (10). The extreme diversity of pachytene piRNAs may allow MIWI and MILI complexes to exert broad effects on the transcriptome through a miRNA-like mechanism.

Summary and Perspective

It is becoming increasingly clear that an ancient and conserved function of the Piwi and piRNA pathway is to protect the genome from the activity of parasitic nucleic acids. Even in ciliates, which diverged earlier than the common ancestor of plants and animals, parallels to the piRNA pathways of flies and mammals are clear. In *Tetrahymena*, the scanning hypothesis for DNA elimination (34) suggests that a complex popu-

lation of small RNAs is first generated from the micronuclear genome and subsequently filtered through interactions with the old macronuclear genome. The small RNAs that emerge from this process specify repeat silencing, in this case by elimination from the newly forming and transcriptionally active macronucleus. DNA elimination depends upon a Piwi protein, *Twil*, but unlike the case in vertebrates and *Drosophila*, also on a Dicer protein (35).

Comparisons to ciliates reveal that, during evolution, the core Piwi and piRNA machinery may have adopted both different strategies for producing and filtering small RNA triggers and different strategies for ultimately silencing targets. In *Drosophila*, the Ping-Pong model strongly suggests a posttranscriptional component to transposon silencing. However there is also evidence for impacts of Piwi proteins on chromatin states (36). In mammals, Piwi proteins have been implicated in DNA methylation (12, 26), a function that may be exerted either directly or indirectly.

Plants lack Piwi proteins and have adapted a different RNAi-based strategy for transposon control. In *Arabidopsis*, the Ago subfamily protein Ago4 is programmed with a complex set of transposon-derived small RNAs (37). In contrast to flies and mammals, in which piRNA loci serve as a genetically encoded reservoir of resistance to mobile elements, each individual transposon copy seems to produce small RNAs in plants. Although the precise mechanisms that funnel expressed repeats into this pathway and exclude protein coding genes have yet to be determined, there are hints that chromatin marks may help to concentrate small RNA production at particular sites. This resembles the situation for centromeric repeats in *S. pombe* where specific histone modifications recruit RNAi components to maintain heterochromatin through a local, self-reinforcing loop of small RNA production (37) that is in many ways analogous to the Ping-Pong amplification loop for piRNAs. Yeast and fly systems differ in their strategies for producing complementary substrates. Where yeast and plants use RNA-dependent RNA polymerases to produce antisense repeat sequences, *Drosophila* and mammals encode them from piRNA loci.

Although much remains to be learned about Piwi proteins and their functions throughout evolution, one must wonder whether this pathway is providing a glimpse into the ancestral functions of RNAi. It seems almost certain that the evolution of genomic parasites followed closely the emergence of self-replicating genomes. Thus, the development of heritable, but also adaptive, systems of parasite resistance would have been essential to maintaining fitness. Given the dire consequences of transposon activation in higher organisms, for example, hybrid dysgenesis in *Drosophila* and sterility in mammals, it is likely that colonization by mobile elements provides a driving force in speciation,

as subpopulations adapt to coexist with specific invaders. The Piwi and piRNA pathway may thus have played a long-standing and important role in maintaining species cohesion by allowing adaptation of populations to new mobile elements and preventing reproductive isolation.

References and Notes

1. N. H. Tolia, L. Joshua-Tor, *Nat. Chem. Biol.* **3**, 36 (2007).
2. X. H. Wang *et al.*, *Science* **312**, 452 (2006).
3. J. G. Ruby *et al.*, *Cell* **127**, 1193 (2006).
4. P. W. Reddien, N. J. Oviedo, J. R. Jennings, J. C. Jenkin, A. Sánchez Alvarado, *Science* **310**, 1327 (2005).
5. J. Brennecke *et al.*, *Cell* **128**, 1089 (2007).
6. K. Saito *et al.*, *Genes Dev.* **20**, 2214 (2006).
7. L. S. Gunawardane *et al.*, *Science* **315**, 1587 (2007).
8. D. N. Cox *et al.*, *Genes Dev.* **12**, 3715 (1998).
9. S. Kuramochi-Miyagawa *et al.*, *Mech. Dev.* **108**, 121 (2001).
10. W. Deng, H. Lin, *Dev. Cell* **2**, 819 (2002).
11. S. Kuramochi-Miyagawa *et al.*, *Development* **131**, 839 (2004).
12. M. A. Carmell *et al.*, *Dev. Cell* **12**, 503 (2007).
13. D. N. Cox, A. Chao, H. Lin, *Development* **127**, 503 (2000).
14. A. Szakmary, D. N. Cox, Z. Wang, H. Lin, *Curr. Biol.* **15**, 171 (2005).
15. A. A. Aravin *et al.*, *Curr. Biol.* **11**, 1017 (2001).
16. V. V. Vagin *et al.*, *Science* **313**, 320 (2006).
17. M. Savitsky, D. Kwon, P. Georgiev, A. Kalmykova, V. Gvozdev, *Genes Dev.* **20**, 345 (2006).
18. A. A. Aravin *et al.*, *Dev. Cell* **5**, 337 (2003).
19. S. Houwing *et al.*, *Cell* **129**, 69 (2007).
20. V. N. Kim, *Genes Dev.* **20**, 1993 (2006).
21. A. Aravin *et al.*, *Nature* **442**, 203 (2006).
22. A. Girard, R. Sachidanandam, G. J. Hannon, M. A. Carmell, *Nature* **442**, 199 (2006).
23. N. C. Lau *et al.*, *Science* **313**, 363 (2006).
24. K. Mochizuki, N. A. Fine, T. Fujisawa, M. A. Gorovsky, *Cell* **110**, 689 (2002).
25. J. V. Hartig, Y. Tomari, K. Forstemann, *Genes Dev.* **21**, 1707 (2007).
26. A. A. Aravin, R. Sachidanandam, A. Girard, K. Fejes-Toth, G. J. Hannon, *Science* **316**, 744 (2007).
27. D. P. Bartel, *Cell* **116**, 281 (2004).
28. N. Prud'homme, M. Gans, M. Masson, C. Terzian, A. Bucheton, *Genetics* **139**, 697 (1995).
29. S. Desset, C. Meignin, B. Dastugue, C. Vaury, *Genetics* **164**, 501 (2003).
30. J. E. Wilson, J. E. Connell, P. M. Macdonald, *Development* **122**, 1631 (1996).
31. C. Klattenhoff *et al.*, *Dev. Cell* **12**, 45 (2007).
32. M. Mevel-Ninio, A. Pelisson, J. Kinder, A. R. Campos, A. Bucheton, *Genetics* **175**, 1615 (2007).
33. D. Bourc'his, T. H. Bestor, *Nature* **431**, 96 (2004).
34. K. Mochizuki, M. A. Gorovsky, *Curr. Opin. Genet. Dev.* **14**, 181 (2004).
35. C. D. Malone, A. M. Anderson, J. A. Motl, C. H. Rexer, D. L. Chalker, *Mol. Cell. Biol.* **25**, 9151 (2005).
36. S. I. Grewal, S. C. Elgin, *Nature* **447**, 399 (2007).
37. R. K. Slotkin, R. Martienssen, *Nat. Rev. Genet.* **8**, 272 (2007).
38. We thank members of the Hannon lab for stimulating discussions. A.A.A. was supported by a Cold Spring Harbor Laboratory Association Fellowship and J.B. was supported by the Schering Foundation. This work was supported by grants from the NIH (G.J.H.) and by a generous gift from Kathryn W. Davis.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/761/DC1
Fig. S1
Table S1

14 June 2007; accepted 17 August 2007
10.1126/science.1146484

Video Cameras on Wild Birds

Christian Rutz,* Lucas A. Bluff, Alex A. S. Weir, Alex Kacelnik

Building on recent advances in microelectronics and communication technologies, we developed miniaturized, animal-borne video cameras for studying the undisturbed behavior of wild, free-ranging birds. Our cameras can record bird behavior in its full ecological, physiological, and social context and in locations and circumstances where conventional observation techniques fail.

New Caledonian crows (*Corvus moneduloides*) are renowned for using tools for extractive foraging (1, 2), but the ecological and evolutionary significances of this unusual behavior are largely unknown. These crows are sensitive to human disturbance and inhabit forested, mountainous areas, where visibility is limited and close following of subjects is impractical. Thus, most current knowledge of the species' biology comes either from observing birds at baited sites or from collecting artifacts (1), but these techniques are not suitable for investigating natural foraging behavior and diets.

We designed tail-mounted camera units (Fig. 1A) that do not interfere with movement and

ensure safe shedding of the tag with regular molt (3). Units transmit a color video signal with sound to custom-built receivers and incorporate very high frequency (VHF) radio tags for simultaneous positional tracking. We deployed cameras on 18 different crows (12 males) in our dry forest study site (21°33'50"S, 165°19'27"E), capturing 451 min of analyzable video footage from 12 subjects (38 ± 5 min per bird, mean \pm SE; maximum of 60 min).

Our footage highlighted the importance of ground foraging for these crows (movie S1). Six birds were terrestrial for 13 to 73% (median of 50%) of analyzable footage (the remaining six were almost exclusively observed in trees), and all video-recorded food items, with the exception of some fruit, were collected on the ground. A case study illustrates how the combination of animal-borne cameras with conventional VHF radio telemetry (video tracking) can generate a rich record of behavior along a known trajectory, yielding data on activity patterns, foraging modes, prey encounters, and diet choice (Fig. 1, B and C, and movie S2).

Video data enabled us to estimate the species' natural foraging efficiency. Our pooled prey sample translates into an encounter rate of eight small food items per hour of ground foraging (118 min of video footage showing birds on the ground across seven subjects with 16 items consumed). Prey items collected during long bouts of ground foraging were considerably smaller than the woodboring beetle larvae (fig. S1) that these crows often hunt with stick tools [Supporting Online Material (SOM) text]. These larvae, a hidden and otherwise unexploited food source, might be an important component of crow diet and could have contributed to the evolution of tool use in this species.

Recordings from two adult males (EK1 and CC1) revealed previously unknown aspects of the species' natural tool use. Crow EK1 used at least three different tools for probing loose substrate on the ground during 45 min of scorable footage (Fig. 1D and movie S3), a foraging mode that has not been reported before (1). The bird (i) used one tool for >18 min, traveling >100 m with it; (ii) briefly put aside this tool on several occasions to use its beak, resuming tool use thereafter; and (iii) transported the tool in flight between prolonged bouts of ground-based tool use and after extracting prey with it. Because sticks are plentiful in forest habitats, these observations indicate that these crows may keep particularly good tools for future use. Both crows made tools on the ground from what appeared to be dry, grasslike stems (movies S2 and S3). This tool material was hitherto unknown (1), and no such tool was found at typical larvae fishing sites in our study area (circa 280 tools collected).

Video tracking is a research tool for studying wild birds that are shy or live in inaccessible habitats and offers three main applications: (i) discovery of unknown behaviors, (ii) context-specific quantification of behavior, and (iii) calibration of other telemetry devices (SOM text, fig. S2, and movie S4).

References and Notes

- G. R. Hunt, R. D. Gray, *Acta Zool. Sinica* 52 (suppl.), 622 (2006).
- L. A. Bluff, A. A. S. Weir, C. Rutz, J. H. Wimpenny, A. Kacelnik, *Comp. Cogn. Behav. Rev.* 2, 1 (2007).
- Materials and methods are available on Science Online.
- We thank N. Barré, S. Blancher, R. Joyce, J. Spaggiari, J. Troscianko, and J. Watts. Supported by the Biotechnology and Biological Sciences Research Council (BB/C517392/1), the Province Sud, the Rhodes Trust (C.R. and L.A.B.), and Linacre College (C.R.). Full acknowledgments are provided in (3).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1146788/DC1
Materials and Methods

SOM Text

Figs. S1 and S2

References and Notes

Movies S1 to S4

20 June 2007; accepted 17 September 2007

Published online 4 October 2007;

10.1126/science.1146788

Include this information when citing this paper.

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

*To whom correspondence should be addressed. E-mail: christian.rutz@zoo.ox.ac.uk

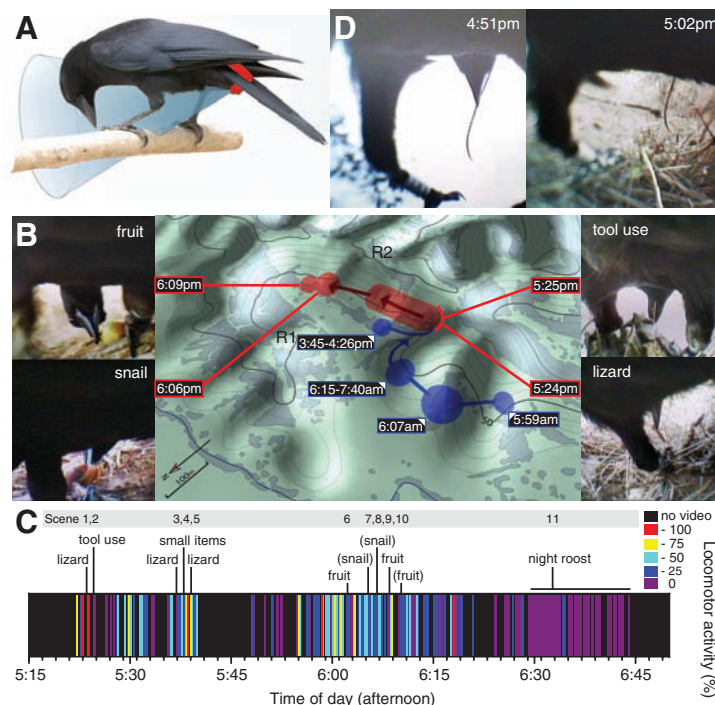


Fig. 1. (A) Position and viewing angle of a video camera (red). The unit is taped to the upper side of two inner tail feathers. The lens at the distal end is protruding through the feathers, peering forward through the legs. (B) Part of a day's foraging trajectory (crow is CC1; 20 December 2006; 5:22 to 6:44 p.m.; movie S2). Live footage was captured by two ground-operated receivers (R1 and R2); a VHF radio tag in the camera unit enabled positional tracking of the bird. Lines show main movements, and oval shapes give approximate locations of the subject before (blue) and during (red) the video shoot. Video-recorded behavior can be linked to the bird's trajectory (see four video stills). Shaded areas indicate shrubs and trees. (C) Activity profile of crow CC1 (percentage locomotor activity in 10-s time bins) during the video shoot ($n = 242$ bins with uninterrupted video reception). Food items in parentheses were not consumed. (D) Tool-assisted ground foraging by a wild crow (EK1; 17 December 2006; 3:56 to 5:06 p.m.; movie S3). Images show a tool (left) during and (right) after a long flight (note distinct curved tip).

Quantum Spin Hall Insulator State in HgTe Quantum Wells

Markus König,¹ Steffen Wiedmann,¹ Christoph Brüne,¹ Andreas Roth,¹ Hartmut Buhmann,¹ Laurens W. Molenkamp,^{1*} Xiao-Liang Qi,² Shou-Cheng Zhang²

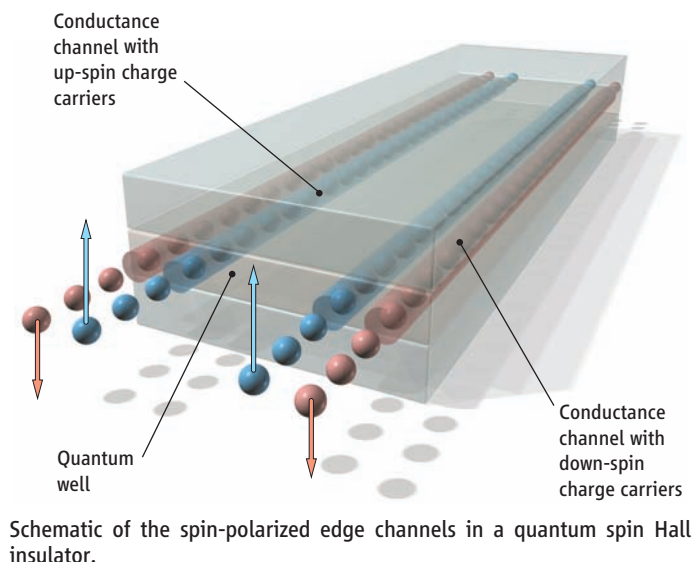
AUTHORS' SUMMARY

The discovery more than 25 years ago of the quantum Hall effect (1), in which the “Hall,” or “transverse electrical” conductance of a material is quantized, came as a total surprise to the physics community. This effect occurs in layered metals at high magnetic fields and results from the formation of conducting one-dimensional channels that develop at the edges of the sample. Each of these edge channels, in which the current moves only in one direction, exhibits a quantized conductance that is characteristic of one-dimensional transport. The number of edge channels in the sample is directly related to the value of the quantum Hall conductance. Moreover, the charge carriers in these channels are very resistant to scattering.

Not only can the quantum Hall effect be observed in macroscopic samples for this reason, but within the channels, charge carriers can be transported without energy dissipation. Therefore, quantum Hall edge channels may be useful for applications in integrated circuit technology, where power dissipation is becoming more and more of a problem as devices become smaller. Of course, there are some formidable obstacles to overcome—the quantum Hall effect only occurs at low temperatures and high magnetic fields.

In the past few years, theoretical physicists have suggested that edge channel transport of current might be possible in the absence of a magnetic field. They predicted (2–4) that in insulators with suitable electronic structure, edge states would develop where—and this is different from the quantum Hall effect—the carriers with opposite spins move in opposite directions on a given edge, as shown schematically in the figure. This is the quantum spin Hall effect, and its observation has been hotly pursued in the field.

Although there are many insulators in nature, most of them do not have the right structural properties to allow the quantum spin Hall effect to be observed. This is where HgTe comes in. Bulk HgTe is a II-VI semiconductor, but has a peculiar electronic structure: In most such materials, the conduction band usually derives from s-states located on the group II atoms, and the valence band from p-states at the VI atoms. In HgTe this order is inverted, however (5). Using molecular beam epitaxy, we can grow thin HgTe quantum wells, sandwiched between (Hg,Cd)Te barriers, that offer a unique way to tune the electronic structure of the material: When the quantum well is wide, the electronic structure in the well remains inverted. However, for narrow wells, it is possible to obtain a “normal” alignment of the quantum well states. Recently, Bernevig *et al.* (6) predicted



theoretically that the electronic structure of inverted HgTe quantum wells exhibits the properties that should enable an observation of the quantum spin Hall insulator state. Our experimental observations confirm this.

These experiments only became possible after the development of quantum wells of sufficiently high carrier mobility, combined with the lithographic techniques needed to pattern the sample. The patterning is especially difficult because of the very high volatility of Hg. Moreover, we have developed a special low-deposition temperature Si-O-N gate insulator (7), which allows us to control the Fermi level (the energy level up to which all

electronics states are filled) in the quantum well from the conduction band, through the insulating gap, and into the valence band. Using both electron beam and optical lithography, we have fabricated simple rectangular structures in various sizes from quantum wells of varying width and measured the conductance as a function of gate voltage.

We observe that samples made from narrow quantum wells with a “normal” electronic structure basically show zero conductance when the Fermi level is inside the gap. Quantum wells with an inverted electronic structure, by contrast, show a conductance close to what is expected for the edge channel transport in a quantum spin Hall insulator. This interpretation is further corroborated by magnetoresistance data. For example, high-magnetic field data on samples with an inverted electronic structure show a very unusual insulator-metal-insulator transition as a function of field, which we demonstrate is a direct consequence of the electronic structure.

The spin-polarized character of the edge channels still needs to be unequivocally demonstrated. For applications of the effect in actual microelectronic technology, this low-temperature effect (we observe it below 10 K) will have to be demonstrated at room temperature, which may be possible in wells with wider gaps.

Summary References

1. K. v. Klitzing, G. Dorda, M. Pepper, *Phys. Rev. Lett.* **45**, 494 (1980).
2. S. Murakami, N. Nagaosa, S.-C. Zhang, *Phys. Rev. Lett.* **93**, 156804 (2004).
3. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
4. B. A. Bernevig, S.-C. Zhang, *Phys. Rev. Lett.* **96**, 106802 (2006).
5. A. Novik *et al.*, *Phys. Rev. B* **72**, 035321 (2005).
6. B. A. Bernevig, T. L. Hughes, S.-C. Zhang, *Science* **314**, 1757 (2006).
7. J. Hinz *et al.*, *Semicond. Sci. Technol.* **21**, 501 (2006).

FULL-LENGTH ARTICLE

Recent theory predicted that the quantum spin Hall effect, a fundamentally new quantum state of matter that exists at zero external magnetic field, may be realized in HgTe/(Hg,Cd)Te quantum wells. We fabricated such sample structures with low density and high mobility in which we could tune, through an external gate voltage, the carrier conduction from n-type to p-type, passing through an insulating regime. For thin quantum wells with well width $d < 6.3$ nanometers, the insulating regime showed the conventional behavior of vanishingly small conductance at low temperature. However, for thicker quantum wells ($d > 6.3$ nanometers), the nominally insulating regime showed a plateau of residual conductance close to $2e^2/h$, where e is the electron charge and h is Planck's constant. The residual conductance was independent of the sample width, indicating that it is caused by edge states. Furthermore, the residual conductance was destroyed by a small external magnetic field. The quantum phase transition at the critical thickness, $d = 6.3$ nanometers, was also independently determined from the magnetic field-induced insulator-to-metal transition. These observations provide experimental evidence of the quantum spin Hall effect.

The theoretical prediction of the intrinsic spin Hall effect in metals and insulators (1–3) has generated great interest in the field of spintronics, because this effect allows direct electric manipulation of the spin degrees of freedom without a magnetic field, and the resulting spin current can flow without dissipation. These properties could lead to promising spintronic devices with low power dissipation.

However, beyond the potential technological applications, the intrinsic spin Hall effect has guided us in the search for new and topologically nontrivial states of matter. The quantum Hall state is the first, and so far the only example, of a topologically nontrivial state of matter, where the quantization of the Hall conductance is protected by a topological invariant. The quantum spin Hall (QSH) insulators (4–6) have a similar, but distinct, nontrivial topological property. The QSH insulators are invariant under time reversal, have a charge excitation gap in the bulk, but have topologically protected gapless edge states that lie inside the bulk insulating gap. This type of insulator is typically realized in spin-orbit coupled systems; the corresponding edge states have a distinct helical property: two states with opposite spin-polarization counterpropagate at a given edge (4, 7, 8). The edge states come in Kramers' doublets, and time-reversal symmetry ensures the crossing of their energy levels at special points in the Brillouin zone. Because of this energy-level crossing, the spectrum of a QSH insulator cannot be adiabatically deformed into that of a topologically trivial insulator without helical edge states; therefore, in this precise sense, the QSH insulators represent a topologically distinct new state of matter.

Theoretical background. It has been proposed theoretically that HgTe/(Hg,Cd)Te quantum wells (QWs) provide a natural re-

alization of the quantum spin Hall effect (6). In zincblende-type semiconductor QWs, there are four relevant bands close to the Fermi level. The $E1$ band consists of the two spin states of the s orbital, whereas the $HH1$ band consists of the $|p_x + ip_y, \uparrow\rangle$ and $|-(p_x - ip_y), \downarrow\rangle$ orbitals. The effective Hamiltonian near the Γ point, the center of the Brillouin zone, is given by

$$H_{\text{eff}}(k_x, k_y) = \begin{pmatrix} H(k) & 0 \\ 0 & H^*(-k) \end{pmatrix},$$

$$H = \varepsilon(k) + d_i(k)\sigma_i \quad (1)$$

where σ_i are the Pauli matrices, and

$$d_1 + id_2 = A(k_x + ik_y) \equiv Ak_+,$$

$$d_3 = M - B(k_x^2 + k_y^2),$$

$$\varepsilon_k = C - D(k_x^2 + k_y^2). \quad (2)$$

Here, k_x and k_y are momenta in the plane of the two-dimensional electron gas (2DEG), and A , B , C , and D are material specific constants. Spin-orbit coupling is naturally built-in in this Hamiltonian through the spin-orbit coupled p orbitals $|p_x + ip_y, \uparrow\rangle$ and $|-(p_x - ip_y), \downarrow\rangle$. Two-dimensional materials can be grouped into three types according to the sign of the Dirac mass parameter M . In conventional semiconductors such as GaAs and CdTe, the s -like $E1$ band lies above the p -like $HH1$ band, and the mass parameter M is positive. Semi-metals such as graphene (9, 10) are described by a massless Dirac model with $M = 0$, although the bands have a different physical interpretation. In so-called inverted semiconductors such as HgTe, the s -like orbital lies below the p -like orbitals; therefore, the Dirac mass parameter M in the HgTe/(Hg,Cd)Te QWs can be continuously tuned from a positive value $M > 0$ for thin QWs with thickness $d < d_c$ to a negative value $M < 0$ for thick QWs with $d > d_c$. A topological quantum phase transition occurs at a critical thickness $d = d_c$, where the system is effectively described by a massless Dirac theory as for graphene. The nature of this

quantum phase transition has also been investigated in more realistic models beyond the simple four-band model presented here, reaching the same conclusion (11).

The QSH phase occurs in the inverted regime where $M < 0$, i.e., when $d > d_c$. The sample edge can be viewed as a domain wall of the mass parameter M , separating the topologically nontrivial phase with $M < 0$ from the topologically trivial phase with $M > 0$, which is adiabatically connected to the vacuum (12). Massless helical states are confined on the sample edge. The sample has a finite conductance even when the Fermi level lies inside the bulk insulating gap. Therefore, as suggested in (6), the QSH state can be experimentally detected by measuring a residual conductance plateau as one varies the gate voltage in the nominally insulating regime. Furthermore, because the current is carried by the edge states, the conductance should be independent of sample width. Protected by the time-reversal symmetry, nonmagnetic impurities or any other time-reversal invariant local perturbations cannot cause elastic back-scattering of the helical edge states, which warrants the topological robustness of the edge state conductance. However, the presence of a magnetic field breaks time-reversal symmetry and therefore can open up a gap in the energy spectrum of the edge states and remove the residual conductance due to the edge states.

Experimental details. We set out to test these theoretical predictions by measuring the transport properties of HgTe/(Hg,Cd)Te QWs as a function of the sample thickness, the gate voltage, and the external magnetic field. We used modulation-doped type III (13) HgTe/Hg_{0.3}Cd_{0.7}Te QW structures fabricated by molecular beam epitaxy (14), with widths (15) varying from 5.5 nm ($d < d_c$) to 12 nm ($d > d_c$). Dedicated low-thermal budget optical and e-beam lithography was used to structure devices in Hall bar geometry with dimensions ($L \times W$) of (600×200) , (20.0×13.3) , (1.0×1.0) , and $(1.0 \times 0.5) \mu\text{m}^2$ (Fig. 1A, inset). All devices had a 110-nm-thick Si₃N₄/SiO₂ multilayer gate insulator (16) and a 5/30 nm Ti/Au gate electrode stack. Transport measurements were done in a ³He/⁴He-dilution refrigerator (base temperature $T < 30$ mK, uniaxial fields up to 18 T) and in a ⁴He cryostat fitted with a vector magnet ($T = 1.4$ K, and fields up to 300 mT in variable direction), using lock-in techniques. At zero gate voltage, the samples were n-type, exhibiting carrier densities between 1.3×10^{11} and $3.5 \times 10^{11} \text{ cm}^{-2}$ and mobilities up to $1.5 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. The carrier density could be reduced continuously by applying a negative gate voltage to the Au electrode with respect to carriers in the QW. The Si-O-N gate insulator stack allowed for quite large gate voltages, enabling us to gate the samples through the gap from n-type to p-type conductance.

¹Physikalisches Institut (EP III), Universität Würzburg, D-97074 Würzburg, Germany. ²Department of Physics, McCullough Building, Stanford University, Stanford, CA 94305-4045, USA.

*To whom correspondence should be addressed. E-mail: molenkmp@physik.uni-wuerzburg.de

Field-induced phase transition. The change in carrier type can be monitored from Hall experiments, as shown in Fig. 1A for a large $[(L \times W) = (600 \times 200) \mu\text{m}^2]$ Hall bar with a well width of 6.5 nm, at 30 mK. The change in carrier type is directly reflected in a sign change of the slope of the Hall resistance R_{xy} , and we can directly infer that the carrier density varies from an electron density $n = 1.2 \times 10^{11} \text{ cm}^{-2}$ at gate voltage $V_g = -1.0 \text{ V}$ to a hole density $p = 1.0 \times 10^{10} \text{ cm}^{-2}$ at $V_g = -2.0 \text{ V}$. At modest magnetic fields, for both n-type and p-type channels, R_{xy} exhibits quantum Hall plateaus, indicative of the good quality of the material, until at fields above $\sim 3 \text{ T}$ the last Landau level is depleted.

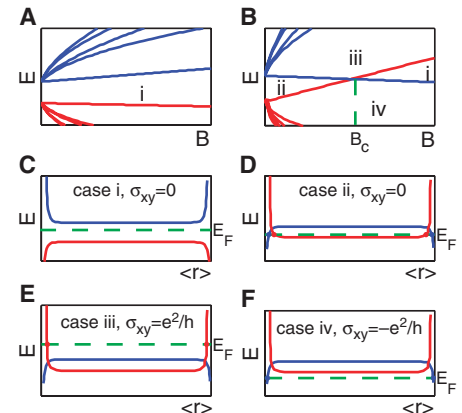
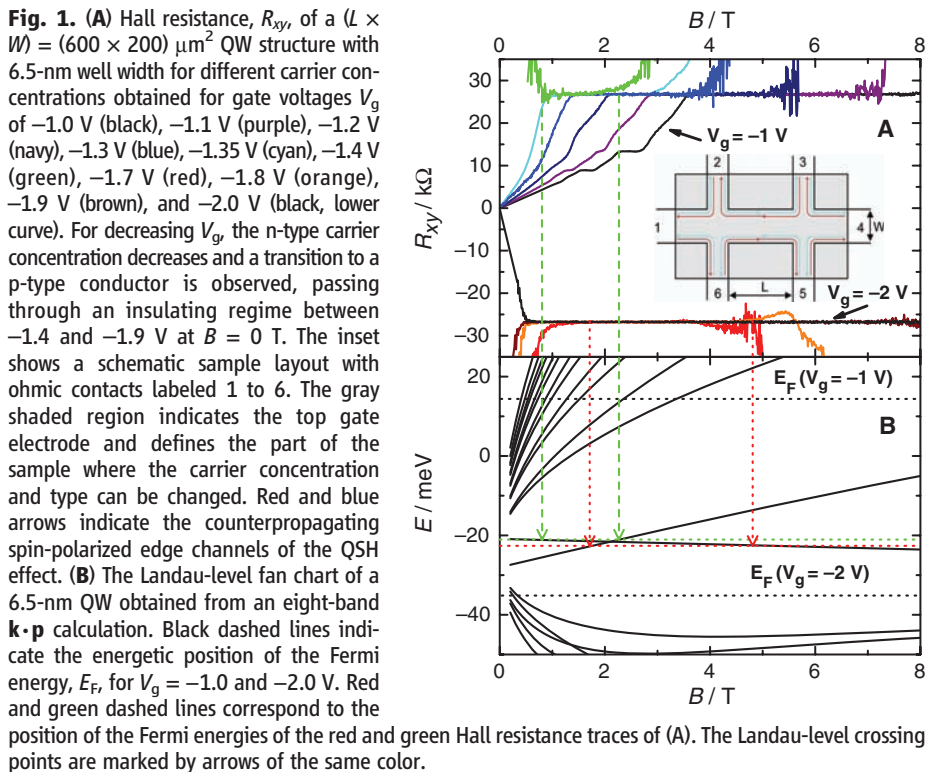
Notable transport behavior is observed for $-1.9 \text{ V} \leq V_g \leq -1.4 \text{ V}$ (see the green and the red traces in Fig. 1), where the sample is insulating at zero magnetic field (i.e., the Fermi level is in the gap). For these gate voltages, we observe that the sample undergoes a phase transition from an insulating state to a QH state with a quantized Hall conductance of $G_{xy} = \pm e^2/h$, either n- or p-type depending on V_g , at a small (~ 1 to 2 T) applied magnetic field. The sample remains in the QH state for a few more T, and then becomes once again insulating. We observed this phenomenon in a number of samples in the inverted regime, with $6.5 \text{ nm} < d < 12 \text{ nm}$.

The phase transition from an insulating state to a QH state is a nontrivial consequence of the inverted band structure of the QSH insulator and can be explained by the level crossing of the $E1$ and $HH1$ Landau levels,

which can be directly obtained from the minimal coupling of the simple Dirac model (Eq. 1) to a perpendicular magnetic field B_\perp . If we only consider the orbital effects of the magnetic field, two series of Landau levels are obtained from the upper and lower 2×2 blocks of the Hamiltonian, Eq. 1, in which the two levels closest to the Fermi energy are given by $E_+ = C + M - (D + B)l_c^{-2}$ and $E_- = C - M - (D - B)l_c^{-2}$, with $l_c \equiv \sqrt{\hbar/(eB_\perp)}$. Thus, the condition for level crossing is given by $E_+ - E_- = 2M - 2Bl_c^{-2} = 0$ or $B_c^\pm = (\hbar M)/(eB)$. Generally, the B parameter is always negative; therefore, we can see that the level crossing occurs only in the inverted region with $M < 0$.

The Landau levels for the normal ($d < d_c$) and the inverted ($d > d_c$) regime are shown in Fig. 2, A and B, respectively. Edge states in the presence of an external magnetic field can be easily obtained by solving our simple Dirac model in Eq. 1 with open boundary condition along one direction and periodic boundary condition along the other direction. Figure 2C shows the bulk and edge states for a conventional insulator. With increasing thickness d , the two states closest to the Fermi energy approach and then cross each other. This “band inversion” leads to the bulk and edge states (Fig. 2D). The Fermi energy crosses a pair of counterpropagating edge states on each edge, resulting in no net Hall effect. These counterpropagating edge states are similar to the helical edge states of the QSH insulator. However, owing to the presence of the magnetic field and the breaking of the time-reversal

symmetry, they are not robustly protected. Starting from this case, increasing the magnetic field shifts the red Landau level toward higher and the blue one toward lower energies. When one of the bulk Landau level crosses the Fermi level, there remains a single edge state on each edge, and one obtains a net QH effect with either $G_{xy} = e^2/h$ (Fig. 2E) or with $G_{xy} = -e^2/h$ (Fig. 2F). When the magnetic field is increased further, the second bulk Landau level crosses the Fermi level, and one reaches the conventional insulator case (Fig. 2C) but with the colors of the Landau level interchanged. In models with bulk inversion asymmetry (BIA) (11), the level crossing between $E1$ and $HH1$ Landau levels at B_c^\pm can be avoided, and the phase regions (i) and (ii) in Fig. 2B become connected. Generally, the nonmonotonic dependence of the Landau-level energies leads to the transition from the insulating state to the QH state at a constant Fermi energy, when the magnetic field is varied.



Although the four-band Dirac model (Eq. 1) gives a simple qualitative understanding of this novel phase transition, we also performed more realistic and self-consistent eight-band $\mathbf{k}\cdot\mathbf{p}$ model calculations (13) for a 6.5-nm quantum well, with the fan chart of the Landau levels displayed in Fig. 1B. The two anomalous Landau levels cross at a critical magnetic field B_{\perp}^c , which evidently depends on well width. This implies that when a sample has its Fermi energy in the gap at zero magnetic field, this energy will always be crossed by the two anomalous Landau levels, resulting in a QH plateau in-between the two crossing fields. Figure 3 summarizes the dependence of B_{\perp}^c on well width d . The open red squares are experimental data points that result from fitting the eight-band $\mathbf{k}\cdot\mathbf{p}$ model to experimental data as in Fig. 1, while the filled red triangles result solely from the $\mathbf{k}\cdot\mathbf{p}$ calculation. For reference, the calculated gap energies are also plotted in this graph as open blue circles. The band inversion is reflected in the sign change of the gap. For relatively wide wells ($d > 8.5$ nm), the (inverted) gap

starts to decrease in magnitude. This is because for these well widths, the band gap no longer occurs between the $E1$ and $HH1$ levels, but rather between $HH1$ and $HH2$ —the second confined hole-like level, as schematically shown in the inset of Fig. 3 [see also (17)]. Also in this regime, a band crossing of conduction- ($HH1$) and valence- ($HH2$) band-derived Landau levels occurs with increasing magnetic field (13, 17, 18). Figure 3 clearly illustrates the quantum phase transition that occurs as a function of d in the HgTe QWs: Only for $d > d_c$ does B_{\perp}^c exist, and at the same time the energy gap is negative (i.e., the band structure is inverted). The experimental data allow for a quite accurate determination of the critical thickness, yielding $d_c = 6.3 \pm 0.1$ nm.

Zero-field edge channels and the QSH effect. The actual existence of edge channels in insulating inverted QWs is only revealed when studying smaller Hall bars [the typical mobility of $10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ in n-type material implies an elastic mean free path of $l_{\text{mfp}} \approx 1 \mu\text{m}$ (19, 20)—and one may anticipate lower

mobilities in the nominally insulating regime]. The pertinent data are shown in Fig. 4, which plots the zero B -field four-terminal resistance $R_{14,23} \equiv V_{23}/I_{14}$ as a function of normalized gate voltage (V_{thr} is defined as the voltage for which the resistance is largest) for several devices that are representative of the large number of structures we investigated. $R_{14,23}$ is measured while the Fermi level in the device is scanned through the gap. In the low-resistance regions at positive $V_g - V_{\text{thr}}$, the sample is n-type; at negative $V_g - V_{\text{thr}}$, the sample is p-type.

The black curve labeled I in Fig. 4 was obtained from a medium-sized $[(20.0 \times 13.3) \mu\text{m}^2]$ device with a 5.5-nm QW and shows the behavior we observe for all devices with a normal band structure: When the Fermi level is in the gap, $R_{14,23}$ increases strongly and is at least several tens of megohm (this is the detection limit of the lock-in equipment used in the experiment). This clearly is the expected behavior for a conventional insulator. However, for all devices containing an inverted QW, the resistance in the insulating regime remains finite. $R_{14,23}$ plateaus at well below 100 kilohm (i.e., $G_{14,23} = 0.3 e^2/h$) for the blue curve labeled II, which is again for a $(20.0 \times 13.3) \mu\text{m}^2$ device fabricated by optical lithography, but that contains a 7.3-nm-wide QW. For much shorter samples ($L = 1.0 \mu\text{m}$, green and red curves III and IV) fabricated from the same wafer, $G_{14,23}$ actually reaches the predicted value close to $2e^2/h$, demonstrating the existence of the QSH insulator state for inverted HgTe QW structures.

Figure 4 includes data on two devices with $d = 7.3$ nm, $L = 1.0 \mu\text{m}$. The green trace (III) is from a device with $W = 1.0 \mu\text{m}$, and the red trace (IV) corresponds to a device with $W = 0.5 \mu\text{m}$. Clearly, the residual resistance of the devices does not depend on the width of the structure, which indicates that the transport occurs through edge channels (21). The traces for the $d = 7.3$ nm, $L = 1.0 \mu\text{m}$ devices do not reach all the way into the p-region because the electron-beam lithography needed to fabricate the devices increases the intrinsic ($V_g = 0$ V) carrier concentration. In addition, fluctuations on the conductance plateaus in traces II, III, and IV are reproducible and do not stem from, e.g., electrical noise. Although all $R_{14,23}$ traces discussed so far were taken at the base temperature (30 mK) of our dilution refrigerator, the conductance plateaus are not limited to this very-low-temperature regime. In the inset of Fig. 4, we reproduce the green 30-mK trace III on a linear scale and compare it with a trace (in black) taken at 1.8 K from another $(L \times W) = (1.0 \times 1.0) \mu\text{m}^2$ sample, which was fabricated from the same wafer. In the fabrication of this sample, we used a lower-illumination dose in the e-beam lithography, resulting in a better (but still not quite complete) coverage of the n-i-p transition. Clearly, in this further sample, and at 1.8 K, the $2e^2/h$ conductance

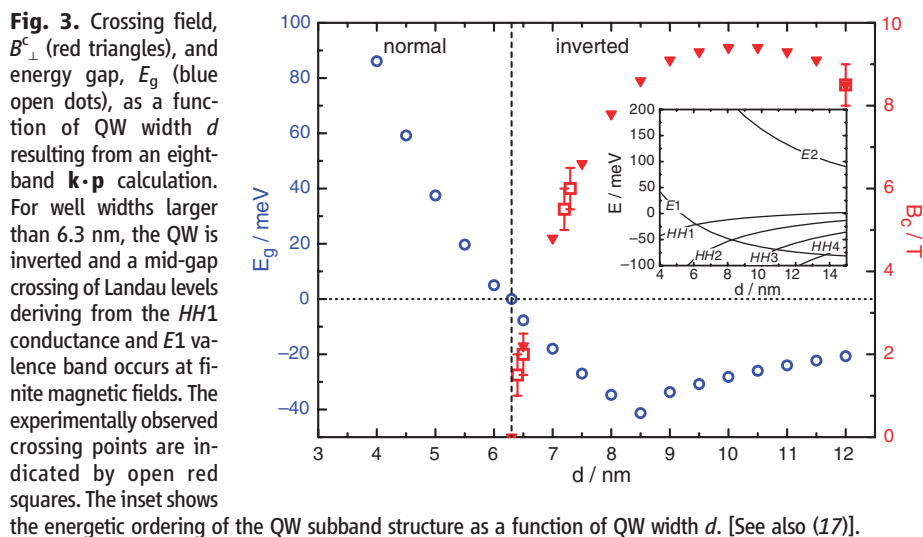


Fig. 3. Crossing field, B_{\perp}^c (red triangles), and energy gap, E_g (blue open dots), as a function of QW width d resulting from an eight-band $\mathbf{k}\cdot\mathbf{p}$ calculation. For well widths larger than 6.3 nm, the QW is inverted and a mid-gap crossing of Landau levels deriving from the $HH1$ conduction and $E1$ valence band occurs at finite magnetic fields. The experimentally observed crossing points are indicated by open red squares. The inset shows the energetic ordering of the QW subband structure as a function of QW width d . [See also (17)].

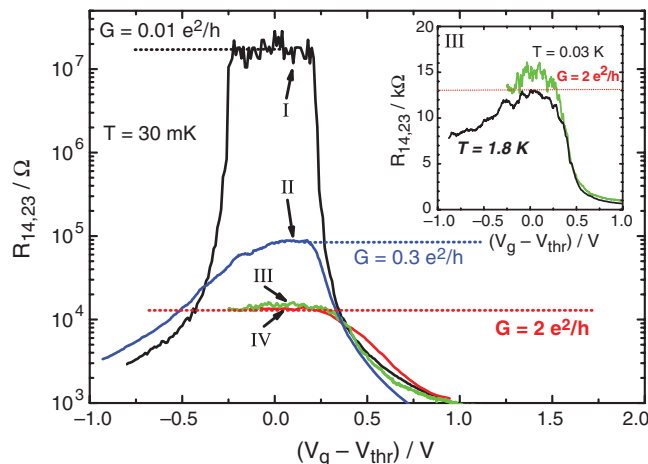


Fig. 4. The longitudinal four-terminal resistance, $R_{14,23}$, of various normal ($d = 5.5$ nm) (I) and inverted ($d = 7.3$ nm) (II, III, and IV) QW structures as a function of the gate voltage measured for $B = 0$ T at $T = 30$ mK. The device sizes are $(20.0 \times 13.3) \mu\text{m}^2$ for devices I and II, $(1.0 \times 1.0) \mu\text{m}^2$ for device III, and $(1.0 \times 0.5) \mu\text{m}^2$ for device IV. The inset shows $R_{14,23}(V_g)$ of two samples from the same wafer, having the same device size (III) at 30 mK (green) and 1.8 K (black) on a linear scale.

plateau is again present, including (thermally smeared) conductance fluctuations.

In the pure two-terminal geometry, with only source and drain contacts (contacts 1 and 4, inset of Fig. 1A), the two counterpropagating helical edge states at one given edge connect the chemical potential from the source and drain, respectively, and they are not in equilibrium with each other because the elastic backscattering vanishes between these two channels. In the absence of voltage probes 2, 3, 5, and 6, as indicated in the inset of Fig. 1A, the two-terminal conductance should give $2e^2/h$. In the presence of the voltage probes, the voltage measurement necessarily leads to the equilibration of the two helical channels with the opposite spin orientation, because the voltage probes are not spin sensitive. A simple Landauer-Büttiker type of calculation shows that the four-terminal resistance should be given by $R_{14,23} = h/2e^2$. In the presence of the voltage probes, the voltage drops V_{12} , V_{23} , and V_{34} add in series to give a higher resistance of $R_{14} \equiv V_{14}/I_{14} = 3h/2e^2$. These results are valid as long as the distance between the voltage probes is less than the inelastic mean free path l_{in} . Although elastic scatterers cannot cause backscattering of the helical edge states, inelastic scatterers can. We estimate the inelastic mean free path to be $l_{in} > 1 \mu\text{m}$ at our measurement temperature. Therefore, for the large sample (trace II), where the distance between the voltage probes exceeds the inelastic mean free path l_{in} , we expect the residual resistance to be higher, consistent with the experimental measurement shown in trace II in Fig. 4.

Breaking time-reversal symmetry. Another intriguing observation is that the QSH effect is destroyed by applying only a small magnetic field perpendicular to the 2DEG plane. Figure 5 shows that the magnetoresistance is strongly anisotropic. (These data were obtained in the vector magnet system at 1.4 K.) A very sharp, cusplike conductance peak is observed for perpendicular field, with the full width at half-maximum (FWHM) $B_{\perp}^{\text{FWHM}} \approx 28 \text{ mT}$ (22).

The peak broadens strongly when the magnetic field is tilted into the QW plane. For fully in-plane fields, the QSH conductance can be observed over a much wider magnetic field range ($B_{\parallel}^{\text{FWHM}} \approx 0.7 \text{ T}$).

The robustness of the helical edge states is ensured by the time-reversal symmetry. A magnetic field breaks time-reversal symmetry and thus turns on a gap between the two otherwise degenerate helical edge states. The perpendicular and in-plane magnetic field lead to different gaps, depending on the matrix elements of the magnetization operator: $E_{\text{gap}\perp} = \langle \uparrow | (\hat{z} \cdot \hat{r} \times \hat{j} + g_{\perp} \mu_B S_{\perp}) | \downarrow \rangle | B |$ and $E_{\text{gap}\parallel} = \langle \uparrow | g_{\parallel} \mu_B S_{\parallel} | \downarrow \rangle | B |$, in which \hat{r}, \hat{j} are the position and electric current operator, respectively, and \hat{z} is the unit-vector perpendicular to the QW plane. $S_{\perp(\parallel)}$ stands for the dimensionless part of the Zeeman-coupling matrix element in a perpendicular (parallel) magnetic field. We can estimate the magnitude of these two gaps by noting that $\langle \uparrow | \hat{z} \cdot \hat{r} \times \hat{j} | \downarrow \rangle \sim e v \xi$ and $\langle \uparrow | S_{\perp(\parallel)} | \downarrow \rangle \sim 1$, in which v and ξ are the Fermi velocity and width of the edge channels, respectively. v and ξ can be obtained from the Dirac parameters as $v = A/\hbar$ and $\xi \approx \hbar v / |M|$. The parameters for the $d = 7.3\text{-nm}$ QW give the dimensionless ratio $e v \xi / \mu_B \sim 280$, which thus leads to $E_{\text{gap}\perp} / E_{\text{gap}\parallel} \sim 10^2$. From this estimate, we can see that the strong anisotropy observed in the experiments originates from the high Fermi velocity of the edge states and the small bulk gap M , which together make the orbital magnetization dominant.

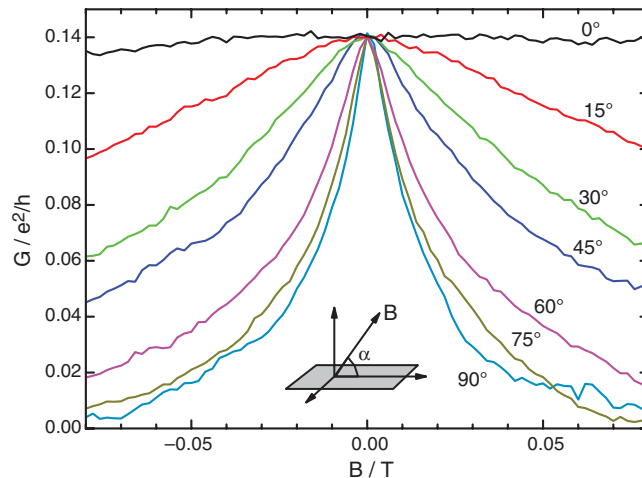
Concluding remarks. Up to this point, our experiments only measured the charge-transport properties. Although the QSH effect was manifested in the change in transport properties, we still want to experimentally confirm the spin accumulation resulting from the spin Hall effect (23, 24) in the topologically nontrivial insulating regime and compare both electric and magnetic results with the experiments of the spin Hall effect in the metallic regime. It would also be interesting to explore the regime close to the quantum phase transition point of $d = d_c$ and compare the transport

properties in that regime with that of the recently discovered graphene. In many ways, the HgTe QW system can be viewed as a tunable graphene system, where the Dirac mass term can be tuned continuously to zero from either the positive (topologically trivial) or the negative (topologically non-trivial) side.

References and Notes

1. S. Murakami, N. Nagaosa, S. C. Zhang, *Science* **301**, 1348 (2003).
2. J. Sinova *et al.*, *Phys. Rev. Lett.* **92**, 126603 (2004).
3. S. Murakami, N. Nagaosa, S. C. Zhang, *Phys. Rev. Lett.* **93**, 156804 (2004).
4. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
5. B. A. Bernevig, S. C. Zhang, *Phys. Rev. Lett.* **96**, 106802 (2006).
6. B. A. Bernevig, T. L. Hughes, S. C. Zhang, *Science* **314**, 1757 (2006).
7. C. Wu, B. A. Bernevig, S. C. Zhang, *Phys. Rev. Lett.* **96**, 106401 (2006).
8. C. Xu, J. Moore, *Phys. Rev. B* **73**, 045322 (2006).
9. K. S. Novoselov *et al.*, *Nature* **438**, 197 (2005).
10. Y. Zhang, Y. Tan, H. L. Stormer, P. Kim, *Nature* **438**, 201 (2005).
11. X. Dai, T. L. Hughes, X.-L. Qi, Z. Fang, S.-C. Zhang, <http://arxiv.org/abs/0705.1516> (2007).
12. Similar mass domain walls have been proposed to occur in three-dimensional in PbTe/(Pb,Sn)Te heterostructures, where the L_z^{\pm} bands change position as a function of Sn concentration (25, 26).
13. E. G. Novik *et al.*, *Phys. Rev. B* **72**, 035321 (2005).
14. C. Becker *et al.*, *Phys. Stat. Sol. (C)* **4**, 3382 (2007).
15. Well thicknesses have been calibrated by x-ray reflectivity measurements at the DESY synchrotron in Hamburg, Germany.
16. J. Hinz *et al.*, *Semicond. Sci. Technol.* **21**, 501 (2006).
17. A. Pfeuffer-Jeschke, thesis, University of Würzburg, Germany (2000).
18. M. Schultz *et al.*, *Phys. Rev. B* **57**, 14772 (1998).
19. V. Daumer *et al.*, *Appl. Phys. Lett.* **83**, 1376 (2003).
20. M. König *et al.*, *Phys. Rev. Lett.* **96**, 076804 (2006).
21. We have observed a similar independence of resistance on sample width in the $d = 7 \text{ nm}$, $L = 20.0 \mu\text{m}$ devices, showing that also in these larger structures, the conductance is completely dominated by edge channels.
22. FWHM of the magnetoresistance peak is about 10 mT at 30 mK, increasing to 28 mT at 1.4 K.
23. Y. K. Kato, R. C. Myers, A. C. Gossard, D. D. Awschalom, *Science* **306**, 1910 (2004).
24. J. Wunderlich, B. Kaestner, J. Sinova, T. Jungwirth, *Phys. Rev. Lett.* **94**, 047204 (2005).
25. B. A. Volkov, O. A. Pankratov, *JETP Lett.* **42**, 178 (1985).
26. E. Fradkin, E. Dagotto, D. Boyanovsky, *Phys. Rev. Lett.* **57**, 2967 (1986).
27. We thank A. Bernevig, X. Dai, Z. Fang, T. Hughes, C. X. Liu, and C. J. Wu for insightful discussions; C. R. Becker and V. Hock for sample preparation; and C. Kumpf for calibrating the well widths of the HgTe samples. This work is supported by the Deutsche Forschungsgemeinschaft (grant SFB 410); the German-Israeli Foundation for Scientific Research and Development (grant 881/05); the NSF (grant DMR-0342832); the U.S. Department of Energy, Office of Basic Energy Sciences, under contract DE-AC03-76SF00515; and Focus Center Research Program (FCRP) Center on Functional Engineered Nanoarchitectonics (FENA).

Fig. 5. Four-terminal magnetoconductance, $G_{14,23}$, in the QSH regime as a function of tilt angle between the plane of the 2DEG and applied magnetic field for a $d = 7.3\text{-nm}$ QW structure with dimensions $(L \times \Omega) = (20 \times 13.3) \mu\text{m}^2$ measured in a vector field cryostat at 1.4 K.



19 June 2007; accepted 10 September 2007

Published online 20 September 2007;

10.1126/science.1148047

Include this information when citing this paper.

Molecular Basis for the Nerve Dependence of Limb Regeneration in an Adult Vertebrate

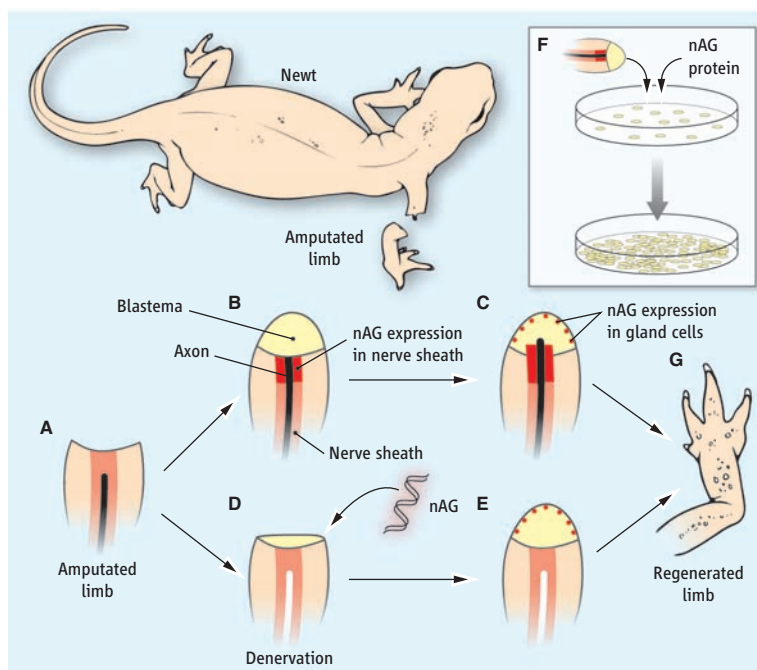
Anoop Kumar,^{1*} James W. Godwin,^{1*} Phillip B. Gates,¹
A. Acely Garza-Garcia,² Jeremy P. Brookes^{1†}

AUTHORS' SUMMARY

Unlike human, salamanders can regrow an amputated limb. This regeneration occurs only if there is simultaneous regeneration of the severed nerves. Although we usually think of the nervous system as carrying information between nerve cells and their sensory and motor targets, nerves are also essential for tissue regeneration. When a salamander limb is amputated at any position, from the shoulder to the fingertips, the stump forms a blastema, a mound of stem cells from which regeneration begins. The nerve supply cut by the amputation also regenerates, and this regrowth is required for the proliferation of the blastemal cells. If the nerves are cut at the base of the limb, deeper than the regenerative tissue, the limb stump is permanently denervated, the axons cannot regenerate, and limb regeneration fails.

The nerve dependence of limb regeneration was discovered in 1823 by Tweedy John Todd, an English physician, and analyzed in the 1940s and 1950s by Marcus Singer. Nerves are required for many different sorts of regeneration in both vertebrates and invertebrates (1), perhaps to ensure that the regenerated tissue receives adequate innervation. Singer showed that either sensory or motor nerves could support regeneration and that neither conduction of the nerve impulse nor neurotransmitter release was required. The molecular basis for communication between nerves and regenerating tissue has been unclear, although various growth factors have been proposed as mediators. Our paper now identifies a protein that can rescue the denervated blastema and induce regeneration of the limb.

The limb blastema gives rise to structures precisely distal to its site of origin (that is, a blastema at the wrist gives rise to a hand, not an upper arm). The stem cells of the blastema can thus be specified according to what structures are being formed, not just what cell types are needed. Analysis of this property led to the previous identification of Prod 1, a small protein anchored to the surface of blastemal cells and expressed in a gradient along the limb (more Prod 1 is found proximally and less at distal points). Prod 1, a member of a protein superfamily called Ly6, determines the proper identity of the regenerating tissues. To search for possible ligands for Prod 1, we conducted a screen in yeast cells and identified the n(ewt)AG protein. This candidate ligand, nAG,



Regeneration of salamander limbs requires concomitant regeneration of the severed nerves. A single protein, nAG, can substitute for the regenerating nerve cells.

showing that expression in both locations depends on axons. In isolated cells cultured from the blastema, nAG promotes proliferation (F). nAG can be artificially expressed in the denervated stump by injecting plasmid DNA that encodes nAG and passing an appropriate current across the limb, a procedure called electroporation that encourages cells to take up external DNA (D). In denervated blastema so treated, nAG is expressed widely and is secreted, and the otherwise missing glands in the wound epidermis are re-established (E). These results suggest that nAG that is normally released from the Schwann cells during regeneration induces the glands (B, C). Most striking, the artificial expression of nAG rescues the denervated blastema so that the limb regenerates through the entire missing proximal-distal axis and forms digits (G). Therefore, the requirement for the presence of nerve cells to achieve normal, regenerated limb shape and structure can be met by expression of a single protein—nAG.

A geneticist might prefer to see the function of nAG tested by leaving the nerve intact and knocking out the expression of the nAG gene in the relevant cells after amputation. This manipulation is not possible on the salamander at present, but the rescue experiment we report may hold promise for future efforts to promote limb regeneration in mammals (2).

Summary References

1. B. M. Carlson, *Principles of Regenerative Biology* (Elsevier, London, 2007).
2. A. Parson, *J. Life Sci.*, p. 60 (September 2007).

is a member of a large conserved family, the anterior gradient proteins, each of which contains a fold named after the protein thioredoxin and is secreted from cells.

With the help of antibodies reactive to two regions of nAG, we identified where in the regenerating newt limb the protein is expressed. After amputation, the severed axons retract within the stump and then grow back along the nerve sheath (see the figure, A and B). The Schwann cells that make up the nerve sheath in this region express nAG (B and C) as the first blastemal cells divide (B). Later, nAG appears within gland cells in the specialized wound epidermis at the end of the limb (C). This epidermis is critical for sustaining later cell proliferation in the blastema. The appearance of nAG in both the Schwann cells and the wound epidermis is abrogated by denervation (D),

FULL-LENGTH ARTICLE

The limb blastemal cells of an adult salamander regenerate the structures distal to the level of amputation, and the surface protein Prod 1 is a critical determinant of their proximodistal identity. The anterior gradient protein family member nAG is a secreted ligand for Prod 1 and a growth factor for cultured newt blastemal cells. nAG is sequentially expressed after amputation in the regenerating nerve and the wound epidermis—the key tissues of the stem cell niche—and its expression in both locations is abrogated by denervation. The local expression of nAG after electroporation is sufficient to rescue a denervated blastema and regenerate the distal structures. Our analysis brings together the positional identity of the blastema and the classical nerve dependence of limb regeneration.

Limb regeneration occurs in various species of salamander and offers important insights into the possibilities for regenerating a complex structure in adult vertebrates (1). Regeneration proceeds from the limb blastema, a mound of mesenchymal stem cells that arises at the end of the stump. A blastema always regenerates structures distal to its level of origin; a wrist blastema gives rise to the hand, whereas a shoulder blastema gives rise to the arm (2). Distal blastemal cells are converted to more proximal cells by exposure to retinoic acid or other retinoids over a relatively narrow range of concentration (3, 4). This finding led to the identification of Prod 1, a determinant of proximodistal (PD) identity that is expressed at the cell surface as a glycosylphosphatidylinositol (GPI)-anchored protein of the Ly6 superfamily (5). Its expression is graded ($P > D$) in both normal and regenerating limbs (6), and distal cells of the larval axolotl blastema are converted to more proximal cells following focal electroporation of a plasmid expressing Prod 1 (7). We have suggested that a ligand for Prod 1 could be an important player in PD identity (5).

The stem cell niche for limb regeneration has been studied intensively, and the key tissues are the regenerating peripheral nerves and the wound epidermis (8). The severed axons retract after amputation and then regenerate back along the nerve sheath and into the blastema. Axonal regeneration can be prevented or arrested by transecting the spinal nerves at the base of the limb, distant from the amputation level (9). The generation of the initial cohort of blastemal cells occurs in a denervated limb, but the growth and division of these cells depends on the concomitant regeneration of peripheral axons (10). Both motor and sensory axons have this activity, and it is independent of impulse traffic or transmitter release (11, 12). If a peripheral nerve is cut and deviated into a skin wound, it can even evoke the formation of an ectopic appendage (13, 14). Limb regeneration is abrogated if the blastema

is denervated during the initial phase of cellular accumulation, but denervation after the mid-bud stage allows the formation of a regenerate (15). The wound epidermis is not required to support proliferation during the first week of regeneration in an adult newt, but it is critical for subsequent division (16).

It remains unclear which molecules are responsible for the activity of the nerve and wound epidermis (8). The candidates considered to date include neuregulin (17, 18), fibroblast growth factor (19), transferrin (20) and substance P (21). In no case has it been demonstrated that a rigorously denervated blastema can be rescued such that it regenerates to form digits. We have identified a secreted protein that is a ligand for Prod 1 and a growth factor for limb blastemal cells. It is induced after amputation as axons regenerate along the nerve sheath, and then appears in the wound epidermis. The expression in both locations is abrogated by denervation. Most notably, the expression of this protein can rescue the denervated limb blastema and support regeneration to the digit stage.

Results

Identification of nAG protein as a ligand of Prod 1. We performed a yeast two-hybrid screen with the 69 amino acid newt Prod 1 protein (without N or C terminal signal/anchoring sequences) as bait and with prey libraries derived from both normal newt limb and limb blastema. In a search for potential extracellular ligands, two secreted proteins were identified as positives from the screen and subsequent control experiments (Fig. 1A). One was a newt member of the family of anterior gradient proteins, originally defined by the XAG2 protein, which is expressed in the cement gland of the *Xenopus* tadpole (Fig. 1B) (22). These proteins have a single thioredoxin fold with a secretory signal sequence (23). They are expressed in secretory epithelia and have been identified as elevated in various examples of rodent and human cancer (24, 25).

Epitope-tagged versions of bacterially expressed nAG and Prod 1 were found to complex together in a standard pull-down assay (Fig. 1C). When myc-tagged nAG was expressed after transfection of mammalian Cos 7 cells, it was secreted and then detected in immunoblots using two different antibodies directed at non-

overlapping sequences (Fig. 1D). The conditioned medium was reacted with live mouse PS cells transfected so as to express the anchored newt Prod 1 on the surface. The binding of nAG to the surface was detected by phosphatase-labeled secondary antibodies (Fig. 1E) and was absent in untransfected PS cells or after reaction with control medium from mock transfected Cos 7 cells.

Nerve-dependent expression in regeneration.

The expression of nAG protein was analyzed by reacting sections of the newt limb with the two antibodies, and these gave comparable results (fig. S1). In the normal limb, there was weak staining of a subset of glands in the dermis (fig. S2), but after amputation the distal end of the nerve sheath reacted strongly, as illustrated by a longitudinal section at 5 days after amputation [post-amputation (pa)], corresponding to the early dedifferentiation stage (26) (Fig. 2A and fig. S3). We analyzed cross sections of the sheath by staining for both nAG and the Schwann cell marker HNK1, and the nAG was expressed in the Schwann cells but not in axons (Fig. 2B). The wound epidermis was initially negative during regeneration, but after day 10 pa it reacted in glandular structures, as shown at day 12 pa, corresponding to the early bud stage (26) (Fig. 2C and fig. S4).

Newts were denervated by cutting the spinal nerves at the brachial plexus of the right limb and then amputated on both sides. The nerve sheath in the innervated limb showed strong expression at day 8 pa, whereas the sheath on the denervated side showed no reactivity (Fig. 3, A and B). Interestingly the expression in the wound epidermis was also dependent on the nerve. Figure 3C shows a low power image of the wound epidermis on the innervated side with nAG positive glands clearly visible, whereas the contralateral limb showed no reactivity and no glandular structures (Fig. 3D). We conclude that the nAG protein is expressed in the key niche tissues early in regeneration and that expression in both locations is abrogated by denervation.

Activities of nAG on the denervated blastema.

To deliver the protein to the adult newt limb, we electroporated plasmid DNA into the distal stump at day 5 pa. In trial experiments, red fluorescent protein (RFP) was strongly expressed in about 30 to 50% of the mesenchymal cells in this region (Fig. 4A) and persisted for up to 3 weeks. We expressed nAG from a plasmid with the N terminal signal sequence, and the protein was readily detectable both in the electroporated cells (Fig. 4B) and after secretion in the extracellular space of the early regenerate. Because this procedure appeared to deliver the protein effectively, we denervated animals on the right side, amputated both limbs, and then electroporated the nAG plasmid or empty vector on the denervated side. At day 8 after electroporation, we sectioned the distal limbs on both sides and stained with the nAG antibodies. None of the animals electroporated with the control vector

¹Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, UK. ²Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: j.brockes@ucl.ac.uk

showed the appearance of nAG positive glands in the wound epidermis, but five out of six animals electroporated with the nAG plasmid showed the induction of nAG positive glands (Fig. 4C). Therefore, the delivery of this protein to a denervated blastema can induce these elements in the wound epidermis.

To determine whether nAG can rescue the nerve dependence of limb regeneration, groups

of animals were denervated on the right side, and then amputated bilaterally (Fig. 4D). At day 5 pa, the right limb was electroporated either with nAG plasmid or with empty vector. The animals were allowed to regenerate, and the progress of limb regeneration was monitored up to day 40 pa. Two representative newts are shown at day 40 in Fig. 4E. The position of the initial denervation is marked with a yellow star.

The newt on the left has regenerated its control left limb, whereas the right denervated limb has failed to regenerate. In some animals, the axons may subsequently regenerate from the level of the star to the amputation plane, but denervated adult newt blastemas undergo fibrosis and other tissue changes that stop them from making a delayed regenerative response (27). All animals electroporated with the vector resembled the left newt in Fig. 4E. The right animal has also regenerated on the control left side, but the expression of nAG has rescued the denervated blastema and regeneration has proceeded to the digit stage. We analyzed the animals of different batches at day 30 to 40 pa, and half of the nAG-electroporated animals showed digit-stage regeneration (Fig. 4E). Some animals regenerated more slowly and were not included as

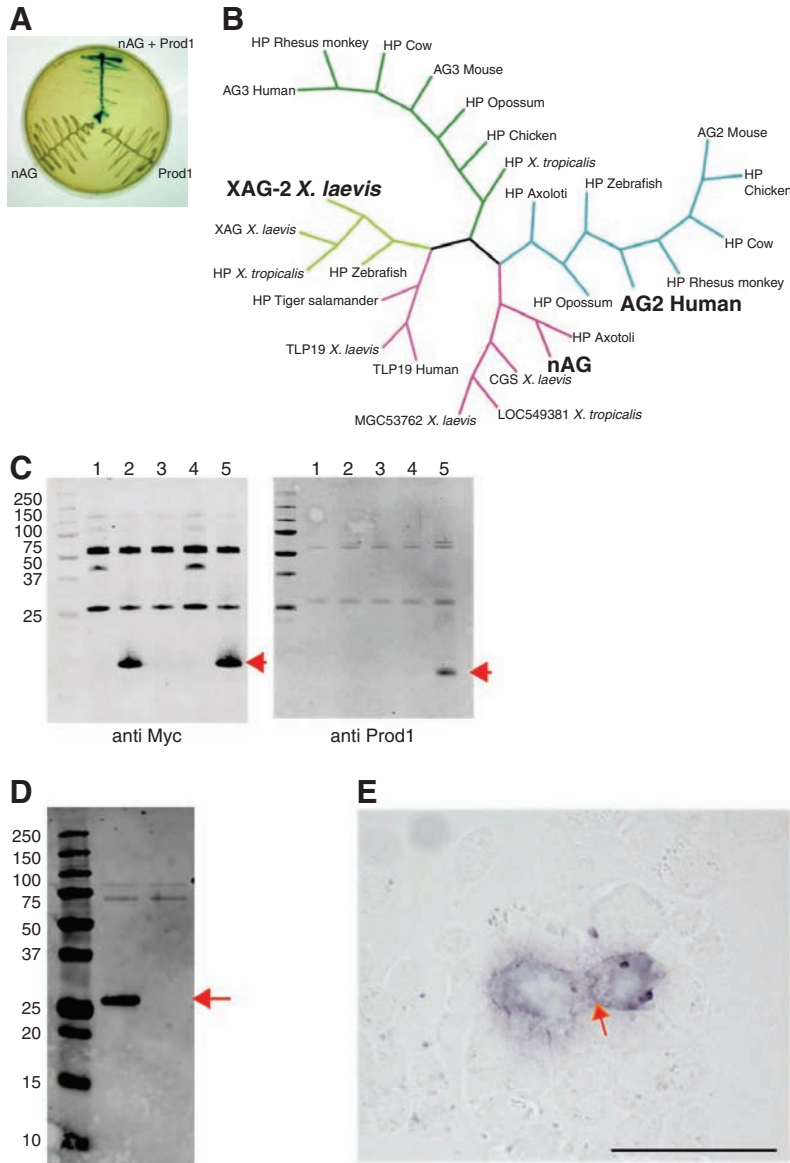


Fig. 1. Identification of nAG protein as a ligand for Prod 1. **(A)** Yeast two-hybrid assay illustrating the interaction between nAG and Prod 1. **(B)** Consensus Bayesian phylogenetic tree of representative members of the AG family of secreted proteins, highlighting nAG, the founder member XAG2, and the human AG2, which is up-regulated in several examples of cancer. **(C)** Pull-down assay with epitope-tagged forms of nAG and Prod 1 purified after bacterial expression. Lane 1, connective tissue growth factor (CTGF) beads; 2, nAG beads; 3, control beads + Prod 1; 4, CTGF beads + Prod 1; 5, nAG beads + Prod 1. Note that Prod 1 is only pulled down in lane 5. **(D)** Secretion of nAG after transfection of Cos 7 cells. Cos 7 cells were transfected with a plasmid expressing the myc-tagged nAG, or RFP, as control. The medium was analyzed by Western blotting with antibody to myc. The central lane is the nAG-transfected sample, the right is the RFP, and the left is the molecular weight markers. **(E)** Reaction of myc-tagged nAG at the surface of Prod 1 transfected mouse PS cells. nAG-conditioned medium derived as in (D) was reacted at 4°C with live PS cells transfected to express Prod 1. Note the purple reaction product at the membrane junction between the two cells (arrow). Scale bar, 50 μm.

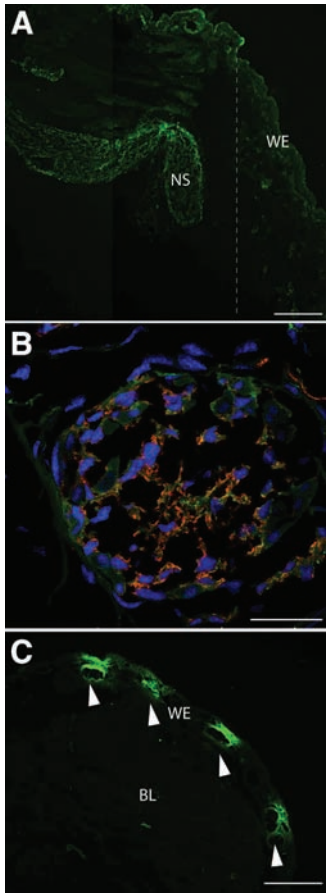


Fig. 2. Expression of nAG after amputation of the adult newt. **(A)** Longitudinal section of a blastema at day 5 pa, stained with antibodies to nAG (green). Note the strong reaction of the nerve sheath and lack of reaction in the WE. The dotted line indicates the position of the amputation plane. **(B)** Cross section of a nerve sheath in a blastema at day 10 pa, stained for nAG (green), the Schwann cell marker HNK1 (red), and nuclei (blue). **(C)** Longitudinal section of a blastema at day 12 pa, stained with antibodies to nAG and showing nAG positive glands (arrow). NS, nerve sheath; BL, blastema; WE, wound epidermis. Scale bars: (A) 200 μm; (B) 50 μm; (C) 250 μm.

reaching digit stage, whereas others did not regenerate, possibly because of the variability in the nAG expression level observed after electroporation of plasmids into adult limbs.

Limbs rescued by nAG expression were sectioned and stained with antibodies, along with

their contralateral control limbs. After staining with antibody to acetylated tubulin, which stains peripheral nerves, the rescued limb showed few labeled profiles, whereas the control limb was densely innervated (Fig. 5, A and B). This result also indicates that nAG

did not rescue the denervated blastema by enhancing the rate of nerve regeneration. These limbs were usually atrophic compared with the contralateral controls (Fig. 4E), and Fig. 5, C and D, shows sections stained with antibody to myosin. The experimental limbs had less muscle than the innervated controls, and it appears that the dependence of skeletal muscle on its innervation (28) was not satisfied by substituting nAG. It is clear, however, that the nerve requirement for completion of the PD axis was met in these animals.

nAG acts as a growth factor for cultured blastemal cells. It is difficult to understand the events underlying cell division in limb mesenchyme because of the complexity of epithelial-mesenchymal interactions in development and regeneration (29, 30). To determine whether nAG acts directly on limb blastemal cells to stimulate their proliferation, the wound epidermis was removed from limb blastemas, and the cells were dissociated and allowed to attach to micro-wells in serum-free medium before maintenance in medium containing 1% serum (Fig. 6A). These cultures were reacted under live conditions with antibody to Prod 1, and ~70% of the cells were specifically stained on their cell surface. The cells were incubated with medium from Cos 7 cells transfected with a nAG plasmid or with a control plasmid. The nAG protein was detected in the medium after immunoblotting under both reducing and nonreducing conditions as a band at 18 kD (fig. S5). The mean stimulation index for S-phase entry, as determined by bromodeoxyuridine (BrdU) pulse labeling, was 8.3 ± 3.3 fold (SD as determined in eight independent experiments) (table S1). All cell preparations were responsive to nAG; an example is shown in Fig. 6, B and C. This evidence supports the view that nAG can rescue the denervated blastema by acting directly on blastemal cells to stimulate their proliferation and, therefore, that it mediates the nerve-dependent growth of the early regenerate.

Discussion

Our identification of nAG as a ligand for the PD determinant Prod 1 has underlined that patterning and cell division are linked at the molecular level. We envisage that PD identity is manifested by the quantitative gradation of Prod 1 (6) and that nAG has no role in specifying that identity but rather acts through Prod 1 to promote cell division. Blastemal growth is stimulated in experimental confrontations of cells differing in positional identity—for example in PD intercalation, in which a wrist level blastema is grafted onto a shoulder stump (31, 32)—and this is always dependent on the presence of the nerve. In a recent study of supernumerary limb formation in the axolotl, the deflection of the brachial nerve into a skin wound provided a growth stimulus to form an ectopic blastema or “bump”; such bumps only progress to form limbs if a piece of skin is grafted from the

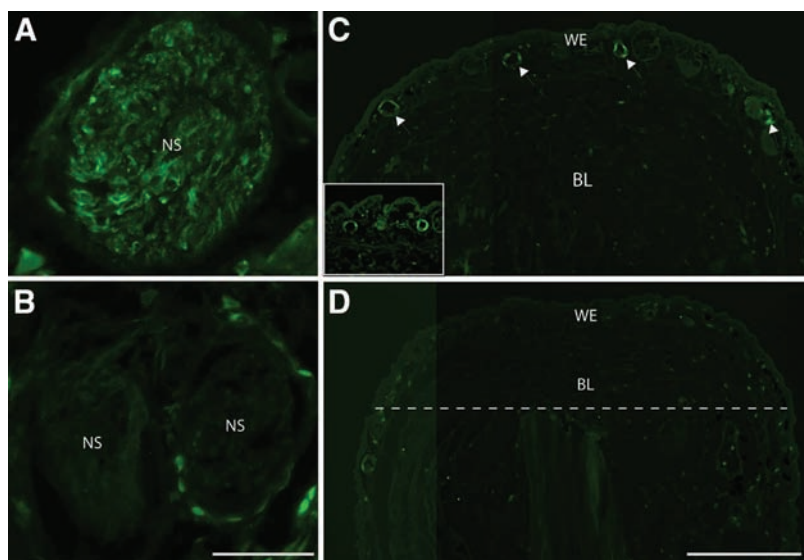
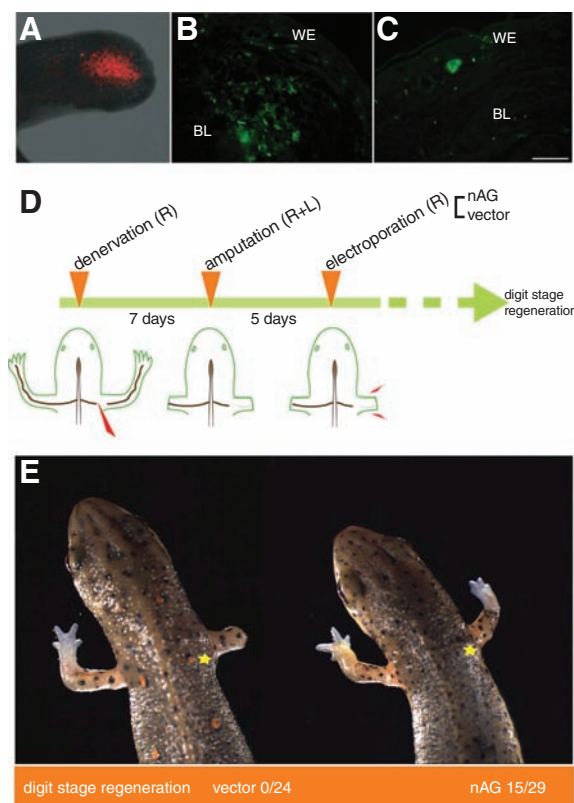


Fig. 3. Expression of nAG in the early blastema depends on innervation. (A) Cross section of a nerve sheath on the innervated side at day 8 pa, and (B) cross section of a sheath on the contralateral denervated side, both stained in parallel with antibodies to nAG. (C) Longitudinal section of a blastema on the innervated side at day 13 pa, showing nAG-positive glands (arrow). The inset shows two glands at higher magnification. (D) Section of the contralateral denervated epidermis, with the amputation plane (dotted line) and the blastema. Scale bars: (A) and (B), 100 μ m; (C) and (D) 500 μ m.

Fig. 4. Delivery of nAG protein to regenerating newt limbs. (A) RFP expression at the end of the limb stump at day 10 pa after electroporation at day 5 pa. (B) Expression of nAG in cells of the limb after electroporation of nAG plasmid at day 7 pa. The section was stained with antibodies to nAG (green). (C) Section of a nAG positive gland in the WE after electroporation of nAG plasmid into a denervated limb blastema at day 5 pa and analysis at day 17 pa. (D) Experimental design for assaying activity of nAG on the denervated blastema. Newts were denervated and amputated before electroporation on the denervated side with either vector or nAG plasmid DNA. (E) Representative animals at day 40 pa from the two groups of an experiment outlined in (D). The yellow star indicates the position of the initial denervation. Scale bars: (B) and (C), 250 μ m.



contralateral skin to the wound site so as to provide dermal fibroblasts of disparate identity (14). It is interesting that dermal fibroblasts express Prod 1 and that this expression is up-regulated by retinoic acid (6).

Two previous studies on AG proteins are relevant to the present results. First, the human AG2 protein was used as bait in a yeast two-hybrid assay and found to complex with a GPI-anchored protein called C4.4, which is associated with metastasis (24, 33). This protein has two Ly6-type domains that are related in sequence to urokinase-type plasminogen activator receptor (34). The degree of relatedness between the three-dimensional structures of Prod 1 and C4.4 domains is not yet resolved, but taken together these results suggest that functional interactions between AG proteins and this class of small Cys-rich protein domains may be conserved.

In the second study, it was found that overexpression of the XAG2 protein in early cleavage stage *Xenopus* embryos could induce formation of an ectopic cement gland that expressed XAG2 (22). We find that expression of nAG induces formation of nAG-positive glands in the denervated newt wound epidermis. After amputation, nAG appears first in the Schwann cells of the distal nerve sheath and then in glands in the wound epidermis. If axonal regeneration is prevented by denervation, neither the Schwann cell nor the glandular expression is detected. Our results suggest that nAG is released by the distal sheath and induces the formation of glands in the wound epidermis. It appears that the secreted nAG acts directly on the limb blastemal cells. It is unclear how the regenerating axons act on the sheath cells, although the membrane form of neuregulin is a candidate, in view of its importance for such interactions (35). The nerve

dependence of regeneration offers a distinction between limb development and regeneration, because the outgrowth of the limb bud is not dependent on its innervation (36). Nonetheless, the ingrowth of the nerve is critical for establishing the nerve dependence, as shown in elegant transplantation experiments on axolotl larvae (37). The identification of nAG offers a new opportunity to study the mechanisms underlying this switch.

Nerve dependence of regeneration is conserved in phylogeny. It has been studied in regeneration of the fish fin, the taste barbel in catfish, the arms of crinoid and asteroid species in echinoderms, and the body axis in annelids (28, 38). In most vertebrate appendages, the density of innervation is lower than in salamanders, and Singer suggested that this is a primary determinant for the loss of regenerative ability, for example in mammals (9, 39). This hypothesis now seems unlikely because there are other variables apparently curtailing regeneration (1). It is notable that the expression of a single protein can rescue limb regeneration in an adult animal (Fig. 4E), and this finding underlines that the blastema is an autonomous unit of organization for which there is no obvious mammalian counterpart. We have suggested that one approach for regenerative medicine would be to understand the specification of the blastema at a level of detail that would allow it to be engineered in mammals (1). The local delivery of permissive regulators such as nAG could then evoke formation of the appropriate structures without the need for subsequent intervention.

Fig. 5. Nerve and skeletal muscle are deficient in nAG-rescued limbs. Rescued limbs were analyzed along with the contralateral innervated limbs, generally at mid-radius/ulna level. Sections of innervated (A) and rescued (B) regenerate limbs were stained with antibody to acetylated tubulin. Sections of innervated (C) and rescued (D) limbs were stained with antibody to skeletal myosin to label muscle. M, muscle. Scale bars, 200 μ m.

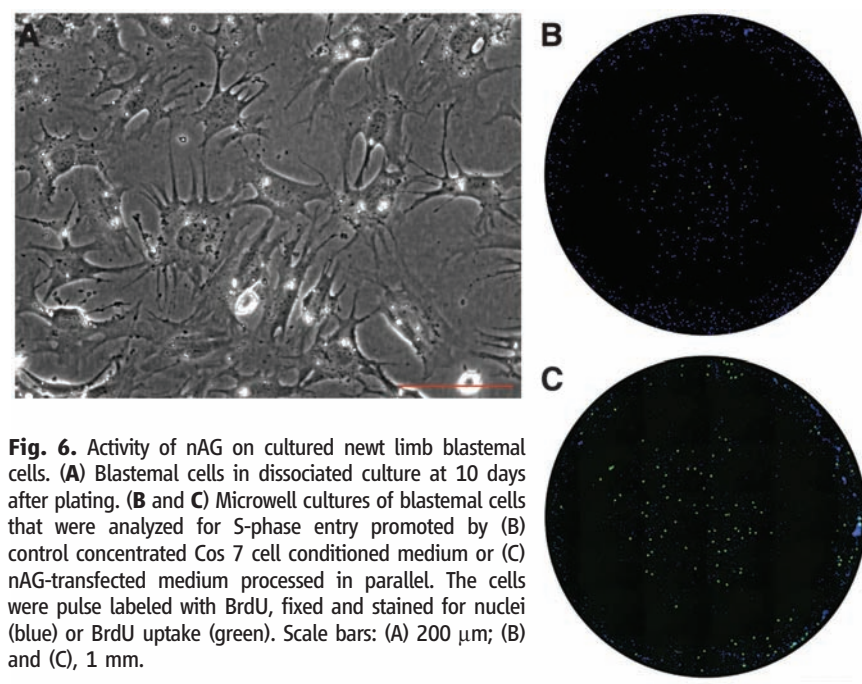
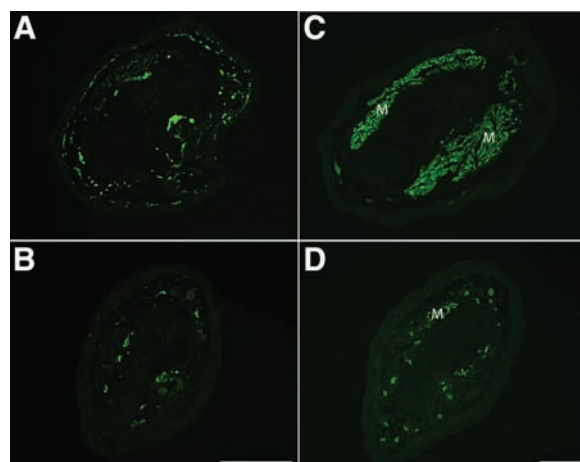


Fig. 6. Activity of nAG on cultured newt limb blastemal cells. (A) Blastemal cells in dissociated culture at 10 days after plating. (B and C) Microwell cultures of blastemal cells that were analyzed for S-phase entry promoted by (B) control concentrated Cos 7 cell conditioned medium or (C) nAG-transfected medium processed in parallel. The cells were pulse labeled with BrdU, fixed and stained for nuclei (blue) or BrdU uptake (green). Scale bars: (A) 200 μ m; (B) and (C), 1 mm.

References and Notes

1. J. P. Brookes, A. Kumar, *Science* **310**, 1919 (2005).
2. J. P. Brookes, *Science* **276**, 81 (1997).
3. M. Maden, *Nature* **295**, 672 (1982).
4. S. D. Thoms, D. L. Stocum, *Dev. Biol.* **103**, 319 (1984).
5. S. Morais da Silva, P. B. Gates, J. P. Brookes, *Dev. Cell* **3**, 547 (2002).
6. A. Kumar, P. B. Gates, J. P. Brookes, *C.R. Biol.* **330**, 485 (2007).
7. K. Echeverri, E. M. Tanaka, *Dev. Biol.* **279**, 391 (2005).
8. D. L. Stocum, *Curr. Top. Microbiol. Immunol.* **280**, 1 (2004).
9. M. Singer, *Q. Rev. Biol.* **27**, 169 (1952).
10. J. P. Brookes, *Science* **225**, 1280 (1984).
11. D. B. Drachman, M. Singer, *Exp. Neurol.* **32**, 1 (1971).
12. R. L. Sidman, M. Singer, *Am. J. Physiol.* **165**, 257 (1951).
13. M. W. Egar, *Anat. Rec.* **221**, 550 (1988).
14. T. Endo, S. V. Bryant, D. M. Gardiner, *Dev. Biol.* **270**, 135 (2004).
15. M. Singer, L. Craven, *J. Exp. Zool.* **108**, 279 (1948).
16. A. L. Mescher, *J. Exp. Zool.* **195**, 117 (1976).
17. J. P. Brookes, C. R. Kintner, *Cell* **45**, 301 (1986).
18. L. Wang, M. A. Marchionni, R. A. Tassava, *J. Neurobiol.* **43**, 150 (2000).
19. L. M. Mullen, S. V. Bryant, M. A. Torok, B. Blumberg, D. M. Gardiner, *Development* **122**, 3487 (1996).
20. A. L. Mescher, E. Connell, C. Hsu, C. Patel, B. Overton, *Dev. Growth Differ.* **39**, 677 (1997).
21. M. Globus, M. J. Smith, S. Vethamany-Globus, *Ann. N. Y. Acad. Sci.* **632**, 396 (1991).
22. F. Aebischer, G. Weidinger, H. Grunz, K. Richter, *Mech. Dev.* **72**, 115 (1998).

23. S. Persson *et al.*, *Mol. Phylogenet. Evol.* **36**, 734 (2005).
24. P. J. Adam *et al.*, *J. Biol. Chem.* **278**, 6482 (2003).
25. D. A. Thompson, R. J. Weigel, *Biochem. Biophys. Res. Commun.* **251**, 111 (1998).
26. L. Iten, S. Bryant, *Willhem Roux Arch. Dev. Biol.* **173**, 263 (1973).
27. J. D. Salley, R. A. Tassava, *J. Exp. Zool.* **215**, 183 (1981).
28. B. M. Carlson, *Principles of Regenerative Biology* (Elsevier Inc., London, 2007).
29. F. V. Mariani, G. R. Martin, *Nature* **423**, 319 (2003).
30. L. Niswander, *Nat. Rev. Genet.* **4**, 133 (2003).
31. K. Crawford, D. L. Stocum, *Development* **102**, 687 (1988).
32. L. T. Pecorino, A. Entwistle, J. P. Brookes, *Curr. Biol.* **6**, 563 (1996).
33. G. C. Fletcher *et al.*, *Br. J. Cancer* **88**, 579 (2003).
34. M. Rosel, C. Claas, S. Seiter, M. Herlevsen, M. Zoller, *Oncogene* **17**, 1989 (1998).
35. G. Lemke, *Sci. STKE* **2006**, pe11 (2006).
36. C. L. Yntema, *J. Exp. Zool.* **142**, 423 (1959).
37. C. S. Thornton, M. T. Thornton, *J. Exp. Zool.* **173**, 293 (1970).
38. R. J. Goss, *Principles of Regeneration* (Academic Press, New York, 1969).
39. M. Singer, K. Rzehak, C. S. Maier, *J. Exp. Zool.* **166**, 89 (1967).
40. We thank P. Driscoll, I. Gout, and P. Martin for their help and comments, and M. Larkum for fabrication of the

electroporation electrodes. This work was supported by a Medical Research Council (UK) Research Professorship and Programme Grant to J.P.B.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/772/DC1
Materials and Methods
Figs. S1 to S5
Table S1
References

12 July 2007; accepted 18 September 2007
10.1126/science.1147710

REPORTS

A Bright Millisecond Radio Burst of Extragalactic Origin

D. R. Lorimer,^{1,2*} M. Bailes,³ M. A. McLaughlin,^{1,2} D. J. Narkevic,¹ F. Crawford⁴

Pulsar surveys offer a rare opportunity to monitor the radio sky for impulsive burst-like events with millisecond durations. We analyzed archival survey data and found a 30-jansky dispersed burst, less than 5 milliseconds in duration, located 3° from the Small Magellanic Cloud. The burst properties argue against a physical association with our Galaxy or the Small Magellanic Cloud. Current models for the free electron content in the universe imply that the burst is less than 1 gigaparsec distant. No further bursts were seen in 90 hours of additional observations, which implies that it was a singular event such as a supernova or coalescence of relativistic objects. Hundreds of similar events could occur every day and, if detected, could serve as cosmological probes.

Transient radio sources are difficult to detect, but they can potentially provide insights into a wide variety of astrophysical phenomena (1). Of particular interest is the detection of short radio bursts, no more than a few milliseconds in duration, that may be produced by exotic events at cosmological distances, such as merging neutron stars (2) or evaporating black holes (3). Pulsar surveys are currently among the few records of the sky with good sensitivity to radio bursts, and they have the necessary temporal and spectral resolution required to unambiguously discriminate between short-duration astrophysical bursts and terrestrial interference. Indeed, they have recently been successfully mined to detect a new galactic population of transients associated with rotating neutron stars (4). The burst we report here, however, has a substantially higher inferred energy output than this class and has not been observed to repeat. This burst therefore represents an entirely new phenomenon.

The burst was discovered during a search of archival data from a 1.4-GHz survey of the Magellanic Clouds (5) using the multibeam receiver on the 64-m Parkes Radio Telescope (6) in Australia. The survey consisted of 209 telescope pointings, each lasting 2.3 hours. During each pointing, the multibeam receiver collected independent signals from 13 different positions (beams) on the sky. The data from each beam were one-bit sampled every millisecond over 96 frequency channels spanning a band 288 MHz wide.

Radio signals from all celestial sources propagate through a cold ionized plasma of free electrons before reaching the telescope. The plasma, which exists within our Galaxy and in extragalactic space, has a refractive index that depends on frequency. As a result, any radio signal of astrophysical origin should exhibit a quadratic shift in its arrival time as a function of frequency, with the only unknown being the integrated column density of free electrons along the line of sight, known as the dispersion measure (DM). Full details of the data reduction procedure to account for this effect, and to search for individual dispersed bursts, are given in the supporting online material. In brief, for each beam, the effects of interstellar dispersion were minimized for 183 trial DMs in the range 0 to 500 cm⁻³ pc. The data were then searched for individual pulses with signal-to-noise (S/N)

ratios greater than 4 with the use of a matched filtering technique (7) optimized for pulse widths in the range 1 to 1000 ms. The burst was detected in data taken on 24 August 2001 with DM = 375 cm⁻³ pc contemporaneously in three neighboring beams (Fig. 1) and was located ~3° south of the center of the Small Magellanic Cloud (SMC).

The pulse exhibited the characteristic quadratic delay as a function of radio frequency (Fig. 2) expected from dispersion by a cold ionized plasma along the line of sight (8). Also evident was a significant evolution of pulse width across the observing frequency band. The behavior we observed, where the pulse width W scales with frequency f as $W \propto f^{-4.8 \pm 0.4}$, is consistent with pulse-width evolution due to interstellar scattering with a Kolmogorov power law [$W \propto f^{-4}$ (9)]. The filter-bank system has finite frequency and time resolution, which effectively sets an upper limit to the intrinsic pulse width $W_{\text{int}} = 5$ ms. We represent this below by the parameter $W_5 = W_{\text{int}}/5$ ms. Note that it is entirely possible that the intrinsic width could be much smaller than observed (i.e., $W_5 \ll 1$) and that the width we observe in Fig. 2 results from the combination of intergalactic scattering and our instrumentation.

We can estimate the flux density of the radio burst in two ways. For the strongest detection, which saturated the single-bit digitizer in the observing system, we make use of the fact that the integrating circuit that sets the mean levels and thresholds is analog. When exposed to a source of strength comparable to the system equivalent flux density, an absorption feature in the profile is induced that can be used to estimate the integrated burst energy. For a 5-ms burst, we estimated the peak flux to be 40 Jy (1 Jy $\equiv 10^{-26}$ W m⁻² Hz⁻¹). Using the detections from the neighboring beam positions, and the measured response of the multibeam system as a function of off-axis position (6), we determined the peak flux density to be at least 20 Jy. We therefore adopt a burst flux of 30 ± 10 Jy, which is consistent with our measurements, for the remaining discussion. Although we have only limited information on the flux density spectrum, as seen in Fig. 2,

¹Department of Physics, West Virginia University, Morgantown, WV 26506, USA. ²National Radio Astronomy Observatory, Green Bank, WV 24944, USA. ³Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia. ⁴Department of Physics and Astronomy, Franklin and Marshall College, Lancaster, PA 17604, USA.

*To whom correspondence should be addressed. E-mail: duncan.lorimer@mail.wvu.edu

the pulse intensity increases at the lowest frequencies of our observing band. This implies that the flux density S scales with observing frequency f as $S \propto f^{-4}$.

It is very difficult to attribute this burst to anything but a celestial source. The frequency dispersion and pulse-width frequency evolution argue for a cosmic origin. It is very unlikely that a swept-frequency transmitter could both mimic the cold plasma dispersion law to high accuracy (see Fig. 2) and have a scattering relation consistent with the Kolmogorov power law. Furthermore, terrestrial interference often repeats, and this was the only dispersed burst detected with $S/N > 10$ in the analysis of data from almost 3000 separate positions. Sources with flux densities greater than ~ 1 Jy are typically detected in multiple receivers of the multibeam system. Although this is true for both terrestrial and astrophysical sources, the telescope had an elevation of $\sim 60^\circ$ at the

time of the observation, making it virtually impossible for ground-based transmitters to be responsible for a source that was only detected in three adjacent beams of the pointing.

We have extensively searched for subsequent radio pulses from this enigmatic source. Including the original detection, there were a total of 27 beams in the survey data that pointed within 30 arcmin of the nominal burst position. These observations, which totaled 50 hours, were carried out between 19 June and 24 July 2001 and showed no significant bursts. In April 2007 we carried out 40 hours of follow-up observations with the Parkes telescope at 1.4 GHz with similar sensitivity to the original observation. No bursts were found in a search over the DM range 0 to $500 \text{ cm}^{-3} \text{ pc}$. These dedicated follow-up observations implied that the event rate must be less than 0.025 hour^{-1} for bursts with $S/N > 6$ (i.e., a 1.4-GHz peak flux density greater than 300 mJy). The data

were also searched for periodic radio signals using standard techniques (8) with null results.

The galactic latitude ($b = -41.8^\circ$) and high DM of the burst make it highly improbable for the source to be located within our Galaxy. The most recent model of the galactic distribution of free electrons (10) predicts a DM contribution of only $25 \text{ cm}^{-3} \text{ pc}$ for this line of sight. In fact, of more than 1700 pulsars currently known, none of the 730 with $|b| > 3.5^\circ$ has $DM > 375 \text{ cm}^{-3} \text{ pc}$. The DM is also far higher than any of the 18 known radio pulsars in the Magellanic Clouds (5), the largest of which is for PSR J0131-7310 in the SMC with $DM = 205 \text{ cm}^{-3} \text{ pc}$. The other four known radio pulsars in the SMC have DMs of 70, 76, 105, and $125 \text{ cm}^{-3} \text{ pc}$. The high DM of PSR J0131-7310 is attributed (5) to its location in an H II region (Fig. 1). We have examined archival survey data to look for ionized structure such as H α filaments or H II regions that could similarly explain the anomalously large DM of the burst. No such features are apparent. The source lies 3° south from the center of the SMC, placing it outside all known contours of radio, infrared, optical, and high-energy emission from the SMC. This and the high DM strongly suggest that the source is well beyond the SMC, which lies $61 \pm 3 \text{ kpc}$ away (11).

No published gamma-ray burst or supernova explosion is known at this epoch or position, and no significant gamma-ray events were detected by the Third Interplanetary Network (12, 13) around the time of the radio burst. The Principal Galaxy Catalog [PGC (14)] was searched for potential hosts to the burst source. The nearest candidate (PGC 246336) is located 5 arcmin south of the nominal burst position, but the nondetection of the burst in the beam south of the brightest detection appears to rule out an association. If the putative host galaxy were similar in type to the Milky Way, the nondetection in the PGC (limiting B magnitude of 18) implies a rough lower limit of $\sim 600 \text{ Mpc}$ on the distance to the source.

We can place an upper bound on the likely distance to the burst from our DM measurement. Assuming a homogeneous intergalactic medium in which all baryons are fully ionized, the intergalactic DM is expected (15, 16) to scale with redshift, z , as $DM \sim 1200 z \text{ cm}^{-3} \text{ pc}$ for $z \leq 2$. Subtracting the expected contribution to the DM from our Galaxy, we infer $z = 0.3$, which corresponds to a distance of $\sim 1 \text{ Gpc}$. This is likely an upper limit, because a host galaxy and local environment could both contribute to the observed DM. Using the electron density model for our Galaxy (10) as a guide, we estimate that there is a 25% probability that the DM contribution from a putative host galaxy is $> 100 \text{ cm}^{-3} \text{ pc}$ and hence $z < 0.2$. Obviously, the more distant the source, the more powerful it becomes as a potential cosmological probe. The sole event, however, offers little hope of a definitive answer at this stage. To enable some

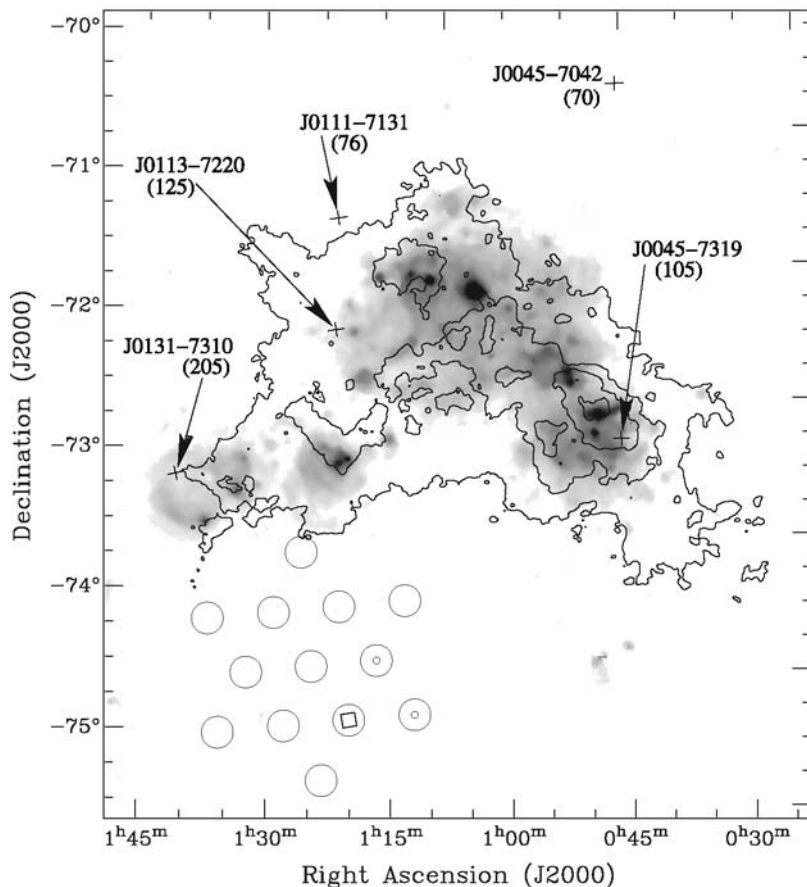


Fig. 1. Multiwavelength image of the field surrounding the burst. The gray scale and contours respectively show H α and H I emission associated with the SMC (32, 33). Crosses mark the positions of the five known radio pulsars in the SMC and are annotated with their names and DMs in parentheses in units of $\text{cm}^{-3} \text{ pc}$. The open circles show the positions of each of the 13 beams in the survey pointing of diameter equal to the half-power width. The strongest detection saturated the single-bit digitizers in the data acquisition system, indicating that its $S/N \gg 23$. Its location is marked with a square at right ascension $01^{\text{h}} 18^{\text{m}} 06^{\text{s}}$ and declination $-75^\circ 12' 19''$ (J2000 coordinates). The other two detections (with S/N s of 14 and 21) are marked with smaller circles. The saturation makes the true position difficult to localize accurately. The positional uncertainty is nominally $\pm 7'$ on the basis of the half-power width of the multibeam system. However, the true position is probably slightly (a few arcmin) northwest of this position, given the nondetection of the burst in the other beams.

indicative calculations about potential source luminosity and event rates, we adopt a distance of 500 Mpc. This corresponds to $z \sim 0.12$ and a host galaxy DM of $200 \text{ cm}^{-3} \text{ pc}$. In recognition of the considerable distance uncertainty, we parameterize this as $D_{500} = D/500 \text{ Mpc}$. If this source is well beyond the local group, it would provide the first definitive limit on the ionized column density of the intracluster medium, which is currently poorly constrained (17).

What is the nature of the burst source? From the observed burst duration, flux density, and distance, we estimate the brightness temperature and energy released to be $\sim 10^{34} (D_{500}/W_5)^2 \text{ K}$ and $\sim 10^{33} W_5 D_{500}^2 \text{ J}$, respectively. These values, and light travel-time arguments that limit the source size to $<1500 \text{ km}$ for a nonrelativistic source, imply a coherent emission process from a compact region. Relativistic sources with bulk velocity v are larger by a factor of either Γ (for a steady jet model) or Γ^2 (for an impulsive blast model), where the Lorentz factor $\Gamma = [1 - (v^2/c^2)]^{-1/2}$ and c is the speed of light.

The only two currently known radio sources capable of producing such bursts are the rotating radio transients (RRATs), thought to be produced by intermittent pulsars (4), and giant pulses from either a millisecond pulsar or a young energetic pulsar. A typical pulse from a RRAT would only be detectable out to $\sim 6 \text{ kpc}$ with our observing system. Even some of the brightest giant pulses from the Crab pulsar, with peak luminosities of 4 kJy kpc^2 (18), would be observable out to $\sim 100 \text{ kpc}$ with the same system. In addition, both the RRAT bursts and giant pulses follow power-law distributions of pulse energies. The strength of this burst, which is some two orders of magnitude above our detection threshold, should have easily led to many events at lower pulse energies, either in the original survey data or follow-up observations. Hence, it appears to represent an entirely new class of radio source.

To estimate the rate of similar events in the radio sky, we note that the survey we have

analyzed was sensitive to bursts of this intensity over an area of about 5 square degrees (i.e., $1/8250$ of the entire sky) at any given time over a 20-day period. Assuming the bursts to be distributed isotropically over the sky, we infer a nominal rate of $8250/20 \approx 400$ similar events per day. Given our observing system parameters, we estimate that a 10^{33}-Jy radio burst would be detectable out to $z \sim 0.3$, or a distance of 1 Gpc. The corresponding cosmological rate for bursts of this energy is therefore $\sim 90 \text{ day}^{-1} \text{ Gpc}^{-3}$. Although considerably uncertain, this is somewhat higher than the corresponding estimates of other astrophysical sources, such as binary neutron star inspirals [$\sim 3 \text{ day}^{-1} \text{ Gpc}^{-3}$ (19)] and gamma-ray bursts [$\sim 4 \text{ day}^{-1} \text{ Gpc}^{-3}$ (20)], but well below the rate of core-collapse supernovae [$\sim 1000 \text{ day}^{-1} \text{ Gpc}^{-3}$ (21)]. Although the implied rate is compatible with gamma-ray bursts, the brightness temperature and radio frequency we observed for this burst are higher than currently discussed mechanisms or limitations for the observation of prompt radio emission from these sources (22).

Regardless of the physical origin of this burst, we predict that existing data from other pulsar surveys with the Parkes multibeam system (23–26) should contain several similar bursts. Their discovery would permit a more reliable estimate of the overall event rate. The only other published survey for radio transients on this time scale (27) did not have sufficient sensitivity to detect similar events at the rate predicted here. At lower frequencies ($\sim 400 \text{ MHz}$) where many pulsar surveys were conducted, although the steep spectral index of the source implies an even higher flux density, the predicted scattering time ($\sim 2 \text{ s}$) would make the bursts difficult to detect over the radiometer noise. At frequencies near 100 MHz , where low-frequency arrays currently under construction will operate (28), the predicted scattering time would be on the order of several minutes, and hence would be undetectable.

Perhaps the most intriguing feature of this burst is its 30-Jy strength. Although this has allowed us to make a convincing case for its extraterrestrial nature, the fact that it is more than 100 times our detection threshold makes its uniqueness puzzling. Often, astronomical sources have a flux distribution that would naturally lead to many burst detections of lower significance; such events are not observed in our data. If, on the other hand, this burst was a rare standard candle, more distant sources would have such large DMs that they would be both red-shifted to lower radio frequencies and outside our attempted dispersion trials. If redshifts of their host galaxies are measurable, the potential of a population of radio bursts at cosmological distances to probe the ionized intergalactic medium (29) is very exciting, especially given the construction of wide-field instruments (30) in preparation for the Square Kilometre Array (31).

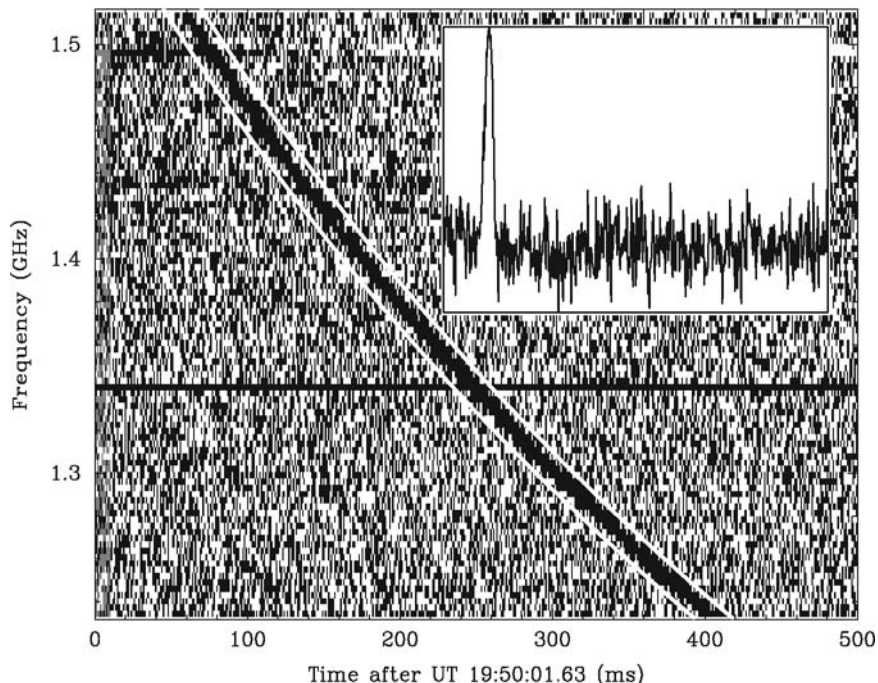


Fig. 2. Frequency evolution and integrated pulse shape of the radio burst. The survey data, collected on 24 August 2001, are shown here as a two-dimensional “waterfall plot” of intensity as a function of radio frequency versus time. The dispersion is clearly seen as a quadratic sweep across the frequency band, with broadening toward lower frequencies. From a measurement of the pulse delay across the receiver band, we used standard pulsar timing techniques and determined the DM to be $375 \pm 1 \text{ cm}^{-3} \text{ pc}$. The two white lines separated by 15 ms that bound the pulse show the expected behavior for the cold-plasma dispersion law assuming a DM of $375 \text{ cm}^{-3} \text{ pc}$. The horizontal line at $\sim 1.34 \text{ GHz}$ is an artifact in the data caused by a malfunctioning frequency channel. This plot is for one of the offset beams in which the digitizers were not saturated. By splitting the data into four frequency subbands, we have measured both the half-power pulse width and flux density spectrum over the observing bandwidth. Accounting for pulse broadening due to known instrumental effects, we determine a frequency scaling relationship for the observed width $W = 4.6 \text{ ms} (f/1.4 \text{ GHz})^{-4.8 \pm 0.4}$, where f is the observing frequency. A power-law fit to the mean flux densities obtained in each subband yields a spectral index of -4 ± 1 . The inset shows the total-power signal after a dispersive delay correction assuming a DM of $375 \text{ cm}^{-3} \text{ pc}$ and a reference frequency of 1.5165 GHz . The time axis on the inner figure also spans the range 0 to 500 ms.

References and Notes

- J. M. Cordes, T. J. W. Lazio, M. A. McLaughlin, *N. Astron. Rev.* **48**, 1459 (2004).
- B. M. S. Hansen, M. Lyutikov, *Mon. Not. R. Astron. Soc.* **322**, 695 (2001).
- M. J. Rees, *Nature* **266**, 333 (1977).
- M. A. McLaughlin *et al.*, *Nature* **439**, 817 (2006).
- R. N. Manchester, G. Fan, A. G. Lyne, V. M. Kaspi, F. Crawford, *Astrophys. J.* **649**, 235 (2006).
- L. Staveley-Smith *et al.*, *Proc. Astron. Soc. Pac.* **13**, 243 (1996).
- J. M. Cordes, M. A. McLaughlin, *Astrophys. J.* **596**, 1142 (2003).
- D. R. Lorimer, M. Kramer, *Handbook of Pulsar Astronomy* (Cambridge Univ. Press, Cambridge, 2005).
- L. C. Lee, J. R. Jokipii, *Astrophys. J.* **206**, 735 (1976).
- J. M. Cordes, T. J. W. Lazio, <http://arxiv.org/abs/astro-ph/0207156> (2002).
- R. W. Hilditch, I. D. Howarth, T. J. Harries, *Mon. Not. R. Astron. Soc.* **357**, 304 (2005).
- K. Hurley *et al.*, *Astrophys. J. Suppl. Ser.* **164**, 124 (2006).
- K. Hurley, personal communication.
- G. Paturel *et al.*, *Astron. Astrophys.* **412**, 45 (2003).
- K. Ioka, *Astrophys. J.* **598**, L79 (2003).
- S. Inoue, *Mon. Not. R. Astron. Soc.* **348**, 999 (2004).
- P. R. Maloney, J. Bland-Hawthorn, *Astrophys. J.* **522**, L81 (1999).
- J. M. Cordes, N. D. R. Bhat, T. H. Hankins, M. A. McLaughlin, J. Kern, *Astrophys. J.* **612**, 375 (2004).
- V. Kalogera *et al.*, *Astrophys. J.* **601**, L179 (2004).
- D. Guetta, M. Della Valle, *Astrophys. J.* **657**, L73 (2007).
- P. Madau, M. Della Valle, N. Panagia, *Mon. Not. R. Astron. Soc.* **297**, L17 (1998).
- J.-P. Macquart, *Astrophys. J.* **658**, L1 (2007).
- R. N. Manchester *et al.*, *Mon. Not. R. Astron. Soc.* **328**, 17 (2001).
- R. T. Edwards, M. Bailes, W. van Straten, M. C. Britton, *Mon. Not. R. Astron. Soc.* **326**, 358 (2001).
- M. Burgay *et al.*, *Mon. Not. R. Astron. Soc.* **368**, 283 (2006).
- B. A. Jacoby, M. Bailes, S. M. Ord, H. S. Knight, A. W. Hotan, *Astrophys. J.* **656**, 408 (2007).
- S. W. Amy, M. I. Large, A. E. Vaughan, *Proc. Astron. Soc. Aust.* **8**, 172 (1989).
- B. W. Stappers, A. G. J. van Leeuwen, M. Kramer, D. Stinebring, J. Hessels, in *Proceedings of the 363. Heraeus Seminar on Neutron Stars and Pulsars*, W. Becker, H. H. Huang, Eds. (Physikzentrum, Bad Honnef, Germany, 2006), pp. 101–103.
- V. L. Ginzburg, *Nature* **246**, 415 (1973).
- S. Johnston *et al.*, *ATNF SKA Memo 13* (Australia Telescope National Facility, 2007).
- P. N. Wilkinson, K. I. Kellermann, R. D. Ekers, J. M. Cordes, T. J. W. Lazio, *N. Astron. Rev.* **48**, 1551 (2004).
- J. E. Gaustad, P. R. McCullough, W. Rosing, D. Van Buren, *Proc. Astron. Soc. Pac.* **113**, 1326 (2001).
- S. Stanimirović, L. Staveley-Smith, J. M. Dickey, R. J. Sault, S. L. Snowden, *Mon. Not. R. Astron. Soc.* **302**, 417 (1999).
- The Parkes Radio Telescope is part of the Australia Telescope, which is funded by the Commonwealth of Australia for operation as a National Facility managed by the Commonwealth Scientific and Industrial Research Organisation. We thank R. Manchester for making the archival data available to us. This research has made use of data obtained from the High Energy Astrophysics Science Archive Research Center, provided by NASA's Goddard Space Flight Center. We thank K. Hurley for providing access to the GCN network archive, and V. Kondratiev, S. Tingay, S. Johnston, F. Camilo, and J. Bland-Hawthorn for useful comments on the manuscript. We acknowledge the prompt awarding of follow-up time by the ATNF Director and thank L. Toomey and P. Sullivan for observing assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1147532/DC1
Materials and Methods
Figs. S1 and S2
References

9 July 2007; accepted 20 September 2007
Published online 27 September 2007;
10.1126/science.1147532
Include this information when citing this paper.

Nanoscale Friction Varied by Isotopic Shifting of Surface Vibrational Frequencies

Rachel J. Cannara,^{1*} Matthew J. Brukman,^{2†} Katherine Cimatu,³ Anirudha V. Sumant,^{2‡} Steven Baldelli,³ Robert W. Carpick^{2§}

Friction converts kinetic energy at sliding interfaces into lattice vibrations, but the detailed mechanisms of this process remain unresolved. Atomic force microscopy measurements reveal that changing the mass of the terminating atoms on a surface, and thus their vibrational frequencies, affects nanoscale friction substantially. We compared hydrogen- and deuterium-terminated single-crystal diamond and silicon surfaces, and in all cases the hydrogenated surface exhibited higher friction. This result implies that the lower natural frequency of chemisorbed deuterium reduces the rate at which the tip's kinetic energy is dissipated. This discovery is consistent with a model describing energy transfer to adsorbates from a moving surface.

Friction converts translational kinetic energy to vibrational energy. Hence, rubbing two bodies together produces heat. This process occurs even in the absence of wear. In contact-mode atomic force microscopy (AFM),

a nanoscale tip slides along a surface, but based on theories of atomic dissipation, some of the tip's translational energy can be converted to lattice vibrations or electronic carriers (1–4). It would be technologically beneficial to control how energy is lost through each of these channels by tuning phononic or electronic properties. Recently, Park *et al.* increased nanoscale friction electronically (5), whereas in the present work we show that friction also depends on the vibrational properties of surfaces.

We altered surfaces by varying the mass, but not the chemistry, of the chemisorbed terminating surface atom. Leaving the surface chemistry unchanged avoids chemical effects due to different interfacial forces. Based on a model of phononic dissipation for friction (6), the surface monolayer acts as an energy-transfer medium, absorbing kinetic energy from the tip

at rates dependent on the adsorbates' natural vibration frequencies (Fig. 1). Because lighter atoms vibrate faster, energy dissipation should be more rapid, and therefore friction should be greater than friction produced by heavier species.

The systems most likely to exhibit observable mass contrast are hydrogen (H)– and deuterium (D)–terminated surfaces, the most durable and inert of which is diamond (7). H- and D-terminated silicon (Si) surfaces are less stable (the surface oxidizes in air after 1 or 2 hours) (8), but studying Si provides an additional test and provides information on an important ma-

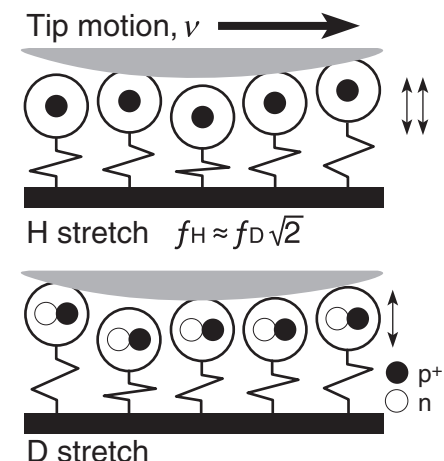


Fig. 1. A schematic of the frictional interface. Vibrating adsorbates collide with and dissipate kinetic energy from the moving tip at a rate that depends on the adsorbate's frequency and thus its mass; that is, at different rates for H than for D.

¹Department of Physics, University of Wisconsin–Madison, Madison, WI 53706, USA. ²Department of Engineering Physics, University of Wisconsin–Madison, Madison, WI 53706, USA. ³Department of Chemistry, University of Houston, Houston, TX 77004, USA.

*Present address: IBM Research GmbH, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland.

†Present address: Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA.

‡Present address: Center for Nanoscale Materials, Argonne National Laboratory, Argonne, IL 60439, USA.

§Present address: Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA 19104, USA.

terial for nanoscale devices and systems. H-terminated diamond also exhibits distinctive electronic properties that depend on the concentration of surface physisorbed contaminants. As described in the supporting online material (9), the absorption of ambient hydrocarbons does indeed change friction for diamond.

The vibrational frequency ω of an adsorbed atom is determined by its mass, the mass of the substrate atom(s) to which it is bonded, and the bond stiffness k . In particular, $\omega = \sqrt{k/\mu}$, where μ is the reduced mass of the bonded adsorbate. In the limit of heavy substrate atoms, μ is approximately the adsorbate's mass, m . On diamond, the ratio of the vibrational frequency of H (ω_H) compared to D (ω_D) is $\omega_H/\omega_D = \sqrt{\mu_D/\mu_H} = 1.363$ (10, 11). On Si(111), the ratio is 1.389 (12–14).

Diamond(001) single-crystal surfaces were terminated with H or D monolayers by means of a hot-filament process (9). This procedure formed well-ordered nearly saturated surfaces, as demonstrated by sum frequency generation (SFG) spectroscopy (Fig. 2A). Vibrational assignments were made on the basis of reported values (15–19). Full coverage of H was achieved, based on the position and intensity of the C-H stretching peak measured at 2835 cm^{-1} on a C(111) surface that was terminated simultaneously with the C(001)-H surface. The peak at 2930 cm^{-1} is assigned to the H-terminated C(001)-(2 \times 1) reconstruction (20). On the deuterated sample, a

single peak at 2160 cm^{-1} was observed, corresponding to the C-D stretch. Analysis of the CH region from 2750 to 3150 cm^{-1} showed no features significantly above the background, demonstrating full coverage by D and a lack of contamination or residual H.

The friction force F_f was measured on diamond as a function of load in both dry nitrogen and ultrahigh vacuum (UHV), alternating multiple times between the H- and D-terminated samples in each case. The same hydrocarbon (HC)-coated tip was used for these experiments. The friction versus load data were fit to a continuum model for the contact area A using the model of interfacial friction $F_f = \tau_0 A$, where τ_0 is the interfacial shear strength (21), a model appropriate in the absence of wear and for load-independent shear strengths. We fit the friction data with a continuum mechanics model describing the dependence of A on load, assuming a spherical, elastic, adhesive contact (22, 23). This model eliminates the need to make any assumptions about the spatial extent of adhesion or the relation between pull-off forces and the work of adhesion.

Caution should be used when applying continuum mechanics to nanoscale contacts, as indicated by recent molecular dynamics simulations (24, 25). Whatever corrections were required, our use of the same tip, substrate, and environment for all H versus D comparisons on diamond, and separately on Si, allowed us to determine unambiguously whether an isotopic substitution effect occurred.

From the continuum mechanics fits, we extracted effective shear strengths (\bar{C}) and work-of-adhesion (γ) values. \bar{C} is proportional to τ_0 , normalized by the appropriate power of the elastic contact modulus (9, 26, 27), which eliminates any dependence on the unknown but constant elastic properties of the tip. Absolute shear strengths (τ_0) were also estimated by selecting reasonable values for the elastic constants of the tip (9), but the effective shear strengths were sufficient for the analysis, because the same tip was always used to compare H- and D-terminated pairs.

The results are summarized in Table 1. Representative plots of \bar{C} for the two terminations and environments studied are shown in Fig. 2B. These data are friction measurements divided by their respective contact area fits at each load, indicating the difference between C-H and C-D beyond the residuals. Deviations at low loads arise from fitting errors intrinsic to the small, steeply changing contact area. The shear strengths (effective and absolute) were an average of 1.26 ± 0.05 and 1.26 ± 0.07 times greater for the H-terminated surfaces for N_2 and UHV, respectively.

Single-crystal Si(111) samples were terminated by H or D by means of an established wet etch process (9), the quality of which was verified by contact angle and x-ray photoelectron spectroscopy measurements. Friction versus load measurements on H- and D-terminated Si(111) samples were obtained in dry N_2 with the same silicon nitride tip. Atomic mass contrast was again observed (Fig. 3 and Table 1), with shear

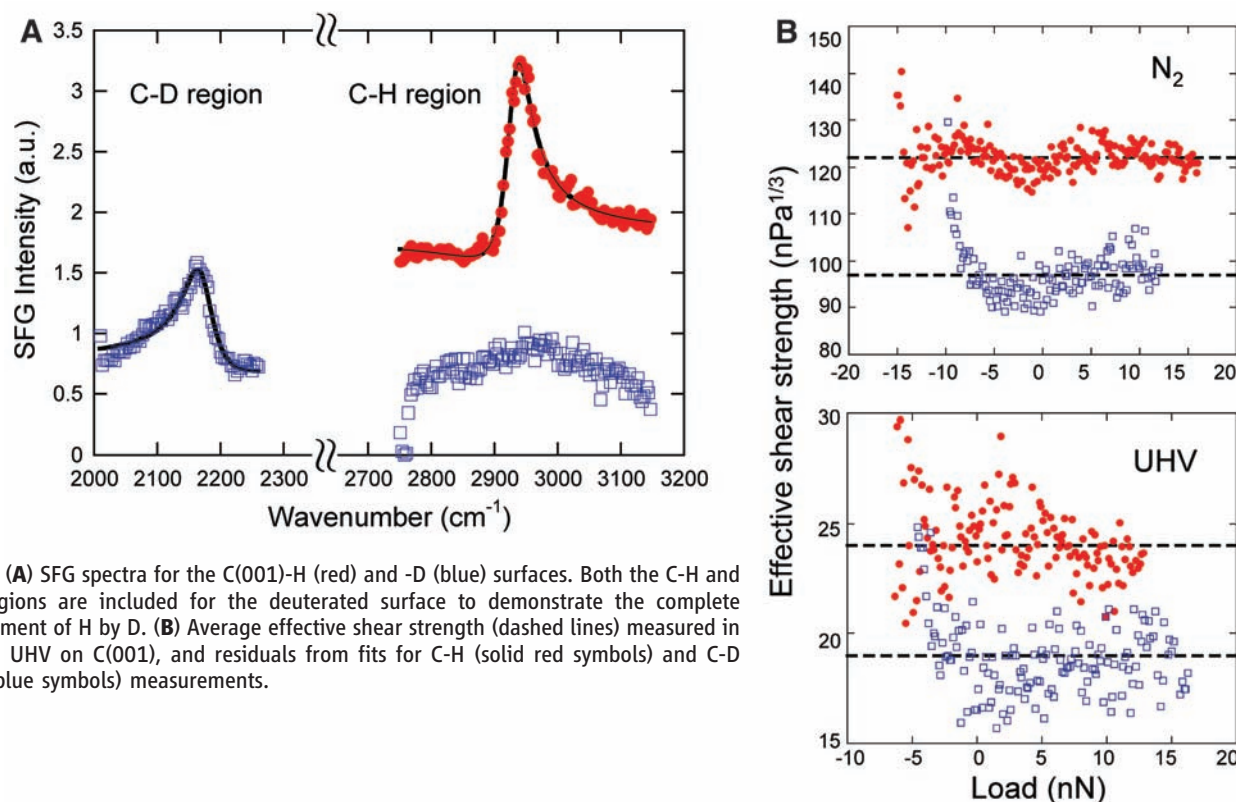


Fig. 2. (A) SFG spectra for the C(001)-H (red) and -D (blue) surfaces. Both the C-H and C-D regions are included for the deuterated surface to demonstrate the complete replacement of H by D. (B) Average effective shear strength (dashed lines) measured in N_2 and UHV on C(001), and residuals from fits for C-H (solid red symbols) and C-D (open blue symbols) measurements.

strengths greater on the H-terminated surface according to the ratio 1.30 ± 0.09 .

The C(001)-H- and -D surfaces have nearly indistinguishable work-of-adhesion values in UHV, consistent with the expected lack of chemical contrast. The work-of-adhesion values are in the range of pure van der Waals interactions (28, 29), particularly for the UHV measurements, and display an impressively small deviation ($<7\%$), which is only expected for well-prepared surfaces in stable environments.

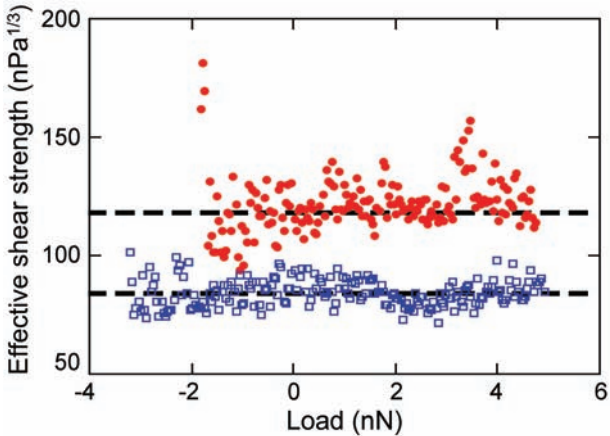
We claim that there is no inherent chemical contrast between the H- and D-terminated surfaces. The small interaction effect from the isotopic substitution of H with D, which can have consequences for very-long-chain polymer systems (30), is calculated to be infinitesimally small here (31). The differences in adhesion that we do observe between H- and D-terminated samples (Table 1) do not vary systematically: Adhesion is higher for H on C(001) in N_2 , lower for H on Si(111) in N_2 , and nearly indistinguishable for C(001) in UHV. The larger variations in N_2 as compared with UHV are probably due to varying degrees of contamination from ambient exposure. We also emphasize that any influence on friction by the contact area (induced by changes in adhesion) has been removed by finding the effective shear strengths.

There are no existing theoretical treatments that permit direct comparison with these experiments.

Table 1. Results for measurements on diamond and Si: the work of adhesion (γ), effective (\tilde{C}) and estimated (τ_0) shear strengths, and their ratio (τ_H/τ_D). The effective shear strength is calculated from the Carpick-Ogletree-Salmeron (COS) transition fits (9, 23). Uncertainties correspond to the standard error on the mean, neglecting systematic uncertainties.

Surface, tip material, environment	Adsorbate	γ (mJ/m ²)	\tilde{C} (nPa ^{1/3})	τ_0 (MPa)	τ_H/τ_D
C(001), HC-coated, N ₂	H	56 ± 2	122 ± 2	1010 ± 17	1.26 ± 0.05
	D	41 ± 2	97 ± 5	806 ± 38	
C(001), HC-coated, UHV	H	30 ± 2	24 ± 1	202 ± 11	1.26 ± 0.07
	D	25 ± 1	19 ± 1	161 ± 7	
Si(111), Si ₃ N ₄ -coated, N ₂	H	113 ± 7	111 ± 7	680 ± 40	1.30 ± 0.09
	D	169 ± 3	86 ± 4	530 ± 20	

Fig. 3. Average effective shear strength (dashed lines) on Si(111) and residuals from representative fits for the Si-H (solid red symbols) and Si-D (open blue symbols) measurements.



However, Persson (6) has modeled phononic friction for an adsorbate monolayer interacting with a single moving surface. Friction arises from inelastic collisions between vibrating adsorbates and the surface. Although this situation does not correspond directly to an AFM measurement where confined atoms are chemisorbed to one surface and interact with the other moving surface under load, we use the model to predict the qualitative behavior. The AFM tip takes the place of the moving surface, and the H (D) atoms correspond to the adsorbates, which are now chemisorbed to the diamond or Si substrate (as opposed to being otherwise unbonded in the model) and are assumed to be uncoupled from one another. In the theory, the friction force from vibrational damping by one adsorbate is

$$F_{f,vib}^{atom} = -m_{tip}\eta v \tag{1}$$

where m_{tip} is the dynamical effective mass of the tip, η is the damping constant of the interaction, and v is the relative sliding velocity between the tip and sample. The minus sign (excluded henceforth) indicates that the friction force opposes the direction of motion. The number of adsorbates involved is proportional to the adsorbate areal density σ and the contact area A . When any variation of η caused by non-uniform contact stresses is neglected, the total vibrational contribution will be $F_{f,vib} = m_{tip}\eta\sigma A$.

For interfacial friction, the vibrational contribution to the shear strength will be

$$\tau_{o,vib} = m_{tip}\eta\upsilon\sigma \tag{2}$$

The damping constant η is related to the density ρ and the transverse sound velocity c_T of the moving surface by

$$\eta \approx 3m\omega^4/8\pi\rho c_T^3 \tag{3}$$

where m and ω refer to the adsorbate’s mass and vibrational frequency, respectively [equations 18 and 22 in (32)].

To illustrate the physics behind Eq. 3, consider an adsorbate vibrating with energy E and colliding with the moving surface (the AFM tip). During one vibration cycle, the adsorbate transfers energy $\Delta E = (m/M_S)E$ from the tip, where M_S is the mass of the portion of the tip effectively involved in the collision (33). Thus, the energy transfer rate may be expressed by $\dot{E} = -\frac{\omega}{2\pi}\left(\frac{m}{M_S}\right)E$. In the collision time $t \approx 2\pi/\omega$, a local displacement field forms in the tip within a distance c_T/ω that corresponds to a volume $(c_T/\omega)^3$. Thus, $M_S \approx \rho(c_T/\omega)^3$. The energy transfer rate is then

$$\dot{E} \approx -\frac{m\omega^4}{2\pi\rho c_T^3}E \tag{4}$$

The solution to Eq. 4 is an exponential decay with damping constant $\eta = (m\omega^4)/(2\pi\rho c_T^3)$, similar to the more precisely derived quantity in Eq. 2 (6). Because $\omega \propto \mu^{-1/2}$, η , τ_0 , and F_f all scale as $m/\mu^2 \approx 1/m$.

It is important to understand how damping leads to friction at the tip/sample interface. Sliding friction is caused by the loss of tip momentum to surface atoms. Atoms excited by the tip can transfer some of their energy back to the tip and help it slip to the next position. This is similar to the concept of reduced friction with increasing temperature (34). [Sørensen *et al.* (35) observed this effect of localized excitations in molecular dynamics simulations.] When the surface potential is symmetric, these excited surface atoms kick the tip in the forward and backward directions, equally promoting and opposing the tip’s motion. Cantilever twist adds a bias to the otherwise symmetric potential and lowers the energy well along the sliding direction (36). Consequently, the momentum transferred back to the tip by surface atoms is more effective in the forward direction, and surface atoms can help push the tip forward. The available energy to do so is reduced by the energy lost from surface atoms to the substrate (that is, damping). Hence, although the rate of excitation vis-à-vis the tip and surface may be governed by the attempt frequencies of the surface atoms, the damping constants represent the diminished ability to transfer energy back to the tip. All of our energy transfer arguments are supported by molecular dynamics simulations, which have

observed H atoms to act as an energy transfer medium to the bulk, specifically for diamond surfaces in sliding contact (37).

This damping model predicts $F_{f,vib}^H/F_{f,vib}^D \approx 1.72$ for diamond and $F_{f,vib}^H/F_{f,vib}^D \approx 1.86$ for Si. Other additive, mass-independent contributions to friction beyond this phononic mechanism will reduce the measured ratios from the theoretical predictions, as observed (Table 1). The predicted ~8% difference in $F_{f,vib}^H/F_{f,vib}^D$ for diamond versus Si is within our experimental uncertainty and may also be counteracted by the increased coupling of surface vibrational modes of deuterated Si(111) to bulk phonons (13) as compared with hydrogenated Si(111) (12). On diamond, surface modes overlap much less with bulk phonons (11). Furthermore, relaxing the assumption that the adsorbates are uncoupled (that is, considering coupled delocalized vibrations) leads to a weaker predicted mass contrast (32). It is likely, however, that surface defects, which accentuate the effect of localized vibrations (9), are present and may even be the dominating contribution to dissipation.

References and Notes

1. M. Cieplak, E. D. Smith, M. O. Robbins, *Science* **265**, 1209 (1994).
2. C. Mak, C. Daly, J. Krim, *Thin Sol. Films* **253**, 190 (1994).
3. B. N. J. Persson, A. I. Volokitin, *J. Chem. Phys.* **103**, 8679 (1995).
4. J. B. Sokoloff, *Phys. Rev. B* **52**, 5318 (1995).

5. J. Y. Park, D. F. Ogletree, P. A. Thiel, M. Salmeron, *Science* **313**, 186 (2006).
6. B. N. J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer, Berlin, ed. 2, 2000), pp. 182–183.
7. T. Ando, M. Ishii, M. Kamo, Y. Sato, *Diam. Rel. Mater.* **4**, 607 (1995).
8. S. Ye *et al.*, *Surf. Sci.* **476**, 121 (2001).
9. Supporting material is available on Science Online.
10. A. Glebov, J. P. Toennies, S. Vollmer, S. A. Safran, J. G. Skofronick, *Phys. Rev. B* **57**, 10082 (1998).
11. S. Thachepan *et al.*, *Phys. Rev. B* **68**, 041401(R) (2003).
12. B. Sandfort, A. Mazur, J. Pollman, *Phys. Rev. B* **51**, 7139 (1995).
13. V. Gräschus, A. Mazur, J. Pollman, *Surf. Sci.* **368**, 179 (1996).
14. C. Stuhlmann, G. Bogdányi, H. Ibach, *Phys. Rev. B* **45**, 6786 (1992).
15. H.-C. Chang, J.-C. Lin, J.-Y. Wu, K.-H. Chen, *J. Phys. Chem.* **99**, 11081 (1995).
16. C. L. Cheng, J. C. Lin, H. C. Chang, J. K. Wang, *J. Chem. Phys.* **105**, 8977 (1996).
17. T. Anzai *et al.*, *J. Mol. Struct.* **352–353**, 455 (1995).
18. R. P. Chin, J. Y. Huang, Y. R. Shen, T. J. Chuang, H. Seki, *Phys. Rev. B* **52**, 5985 (1995).
19. R. P. Chin *et al.*, *Phys. Rev. B* **45**, 1522 (1992).
20. L. V. Zhigilei, D. Srivastava, B. J. Garrison, *Surf. Sci.* **374**, 333 (1997).
21. M. Enachescu *et al.*, *Tribol. Lett.* **7**, 73 (1999).
22. D. S. Grierson, E. E. Flater, R. W. Carpick, *J. Adh. Sci. Tech.* **19**, 291 (2005).
23. R. W. Carpick, D. F. Ogletree, M. Salmeron, *J. Colloid Interf. Sci.* **211**, 395 (1999).
24. B. Luan, M. O. Robbins, *Nature* **435**, 929 (2005).
25. B. Luan, M. O. Robbins, *Phys. Rev. E* **74**, 26111 (2006).
26. U. D. Schwarz, O. Zworner, P. Koster, R. Wiesendanger, *Phys. Rev. B* **56**, 6987 (1997).
27. G. Gao, R. J. Cannara, R. W. Carpick, J. A. Harrison, *Langmuir* **23**, 5394 (2007).
28. J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, London, ed. 2, 1992).
29. A. V. Sumant *et al.*, *Adv. Mat.* **17**, 1039 (2005).
30. F. S. Bates, G. D. Wignall, W. C. Koehler, *Phys. Rev. Lett.* **55**, 2425 (1985).
31. M. J. Brukman, thesis, University of Wisconsin–Madison, Madison, WI (2005).
32. B. N. J. Persson, E. Tosatti, D. Fuhrmann, G. Witte, C. Woll, *Phys. Rev. B* **59**, 11777 (1999).
33. R. E. Walkup, D. M. News, P. Avouris, in *Atomic and Nanometer-Scale Modification of Materials: Fundamentals and Applications*, P. Avouris, Ed. (Kluwer, Dordrecht, Netherlands, 1993), p. 100.
34. X. Zhao, M. Hamilton, W. G. Sawyer, S. S. Perry, *Tribol. Lett.* **27**, 113 (2007).
35. M. R. Sørensen, K. W. Jacobsen, P. Stoltze, *Phys. Rev. B* **53**, 2101 (1996).
36. G. A. Tomlinson, *Philos. Mag.* **7**, 905 (1929).
37. J. A. Harrison, C. T. White, R. J. Colton, D. W. Brenner, *Thin Solid Films* **260**, 205 (1995).
38. We thank J. Butler and J. Yang for loaning and polishing the (001)-oriented single-crystal diamonds and for helpful advice regarding H termination. J. Harris supplied (111) single-crystal diamonds used as references in the SFG measurements. We also thank L. Bruch for extensive discussions and B. Gotsmann for a helpful reading of this manuscript. Supported by the NSF CAREER Award no. CMS 0134571; an NSF Graduate Research Fellowship; and the Air Force Office of Scientific Research, grant FA9550-05-1-0204.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/780/DC1
Materials and Methods
SOM Text
Fig. S1
References

9 July 2007; accepted 21 September 2007
10.1126/science.1147550

A Predictably Selective Aliphatic C–H Oxidation Reaction for Complex Molecule Synthesis

Mark S. Chen and M. Christina White*

Realizing the extraordinary potential of unactivated sp^3 C–H bond oxidation in organic synthesis requires the discovery of catalysts that are both highly reactive and predictably selective. We report an iron (Fe)-based small molecule catalyst that uses hydrogen peroxide (H_2O_2) to oxidize a broad range of substrates. Predictable selectivity is achieved solely on the basis of the electronic and steric properties of the C–H bonds, without the need for directing groups. Additionally, carboxylate directing groups may be used to furnish five-membered ring lactone products. We demonstrate that these three modes of selectivity enable the predictable oxidation of complex natural products and their derivatives at specific C–H bonds with preparatively useful yields. This type of general and predictable reactivity stands to enable aliphatic C–H oxidation as a method for streamlining complex molecule synthesis.

The 20th century witnessed tremendous advances in synthetic methods and strategies that have enabled small molecule targets of extraordinary complexity and biological importance to be synthesized in the laboratory (1). An important remaining challenge is to achieve syntheses with heightened levels of

efficiency. Because many biologically relevant small molecules are oxidized hydrocarbons, reactions that incorporate oxidized functionality selectively into organic frameworks are of particular interest in this regard. Three general reaction classes have been developed for this purpose: functional group interconversions, C–C bond-forming reactions of preoxidized fragments, and olefin oxidations. With these reactions, modern synthetic planning often centers around the use and maintenance of preexisting oxidized functionality. A powerful new class of reactions is

emerging that introduce oxidized functionality directly into aliphatic (sp^3) C–H bonds. Oxidation reactions for isolated, unactivated sp^3 C–H bonds capable of operating with predictable selectivities on complex substrates hold special promise for streamlining syntheses. Such reactions would provide a general way to install oxidized functionalities at a late stage, thereby reducing unproductive chemical manipulations associated with carrying them through a sequence (2, 3).

Despite important advances in the discovery of catalytic methods for aliphatic C–H bond hydroxylations, aminations, and alkylations (4–6), selective reactivity with complex substrates has only been demonstrated for activated C–H bonds (i.e., adjacent to a heteroatom or π system) (7–11) or via the use of substrate directing groups (12–14). High-yielding oxidations of isolated, unactivated sp^3 C–H bonds are rare, and predictable reactivity has only been shown with simple hydrocarbon substrates (10, 15–17). The paradoxical challenge in solving this problem lies in discovering a catalyst that is both highly reactive and predictably selective for oxidizing these inert and ubiquitous C–H bonds. Moreover, to be useful in complex molecule synthesis, this reactivity and selectivity must be general for a broad range of densely functionalized substrates. Nature's design principles for creating such catalysts involve the use of elaborate protein binding pockets that inherently limit substrate generality. A different

Department of Chemistry, Roger Adams Laboratory, University of Illinois, Urbana, IL 61801, USA.

*To whom correspondence should be addressed. E-mail: white@scs.uiuc.edu

strategy was suggested to us by seminal work on site-selective olefin oxidations using bulky, electrophilic metal catalysts (18, 19). With these reagents, mono-oxidation of polyenes occurs predictably at the most electron-rich, least sterically hindered double bond. Moreover, polar functionality proximal to the olefin can direct oxidation, overriding electronic and steric effects. We hypothesized that site-selective oxidations of unactivated sp^3 C–H bonds could similarly be predictably controlled if a suitably reactive metal catalyst could be discovered that is capable of discriminating the subtle electronic and steric differences between C–H bonds in complex molecules (Fig. 1). We herein report an electrophilic iron catalyst, **4**, with a bulky ligand framework that uses H_2O_2 , an inexpensive, environmentally friendly oxidant to effect highly selective oxidations of unactivated sp^3 C–H bonds over a broad range of substrates. We demonstrate that the site of oxidation with **4** can be predicted in complex organic substrates on the basis of the electronic and steric environment of the C–H bond. Additionally, when carboxylate functionality is present, it can direct oxidations toward five-membered ring lactone formation.

Several nonheme iron complexes have shown promising, stereospecific hydroxylation reactivities with unactivated sp^3 C–H bonds in simple hydrocarbon substrates (20–23). The application of these systems to preparative C–H oxidation chemistry has been prevented by the requirement for large excesses of substrate relative to oxidant, low catalyst turnover numbers, and poor selectivities for product formation. Nevertheless, iron(mep) complexes [mep is *N,N'*-dimethyl-*N,N'*-bis(2-pyridylmethyl)-ethane, 1,2-diamine] appeared promising for preparative C–H oxidations with complex substrates because they operate via an electrophilic metal oxidant (22, 23), have a bulky ligand framework amenable to modification, and have been used for preparative epoxidations of olefins containing functionality (24).

The attempted C–H oxidation of pivalate **1** with electrophilic $[Fe(II)(mep)(CH_3CN)_2](SbF_6)_2$ complex, **3**, under preparative conditions (substrate as the limiting reagent) resulted in only low conversion of starting material (12%) and modest selectivity for formation of tertiary hydroxylated product, **2** (56%, Fig. 2A, entry 1). Previous studies have shown a positive correlation between flexibility of the mep ligand and the lability of its iron complexes under oxidative conditions (25). Because unselective oxidations with nonheme iron complexes are often attributed to Fenton-type chemistry upon catalyst decomposition (25), we hypothesized that increasing the rigidity of the mep ligand may lead to improved selectivities. Exchanging the ethylene bridge with a cyclohexane ring had no effect. However, incorporating the methylamines into rigidifying pyrrolidine rings, which furnishes crystallographically characterized com-

plex **4** (26), showed a notable improvement in selectivity (92%), translating into a doubling of the yield of **2** (14% yield, entry 2). The addition of acetic acid (AcOH), previously demonstrated to have a beneficial effect on epoxidations with **3** (24), increased the catalytic activities of both **3** and **4** without substantially changing their selectivities (entries 3 and 4, respectively). This resulted in notable improvement in yields with catalyst **4** (entry 4). Whereas increasing initial catalyst loadings, equivalents of AcOH, or equivalents of H_2O_2 (alone or in combination) gave no further improvements in yield, collective addition of all three components in a portionwise manner furnished preparatively useful amounts of hydroxylated product. Specifically, we found that three consecutive additions of catalyst **4** (5 mol%), AcOH (0.5 equivalent), and H_2O_2 (1.2 equivalents) over a period of 30 min afforded

diastereomerically pure hydroxylated product **2** in 51% isolated yield (entry 5).

A preliminary investigation of the substrate scope highlights the selective, electrophilic nature of the oxidant generated with **4** and H_2O_2 (Fig. 3). In all cases examined, hydroxylation occurred preferentially at the most electron-rich tertiary (3°) C–H bond, despite the fact that secondary (2°) C–H bonds have a significant statistical advantage (entries 1 to 9). Although the dicationic iron catalyst **4** is Lewis acidic, a remarkable range of moderately Lewis basic groups were well tolerated. For example, cyclic ethers, esters, carbonates, and electron-deficient amides were compatible with this C–H oxidation reaction (Fig. 3). In all cases examined in which the 3° C–H bond is part of a stereogenic center, hydroxylation occurred with complete retention of stereochemistry (entries 6 and 7).

Fig. 1. Comparison between established modes of site-selective olefin oxidations and proposed modes for site-selective C–H bond oxidations with bulky, electrophilic metal catalysts. BG indicates bulky group; DG, directing group. (I) In asymmetric dihydroxylations (AD) catalyzed by electrophilic OsO_4 with bulky quinuclidine ligands, dihydroxylation of polyenes occurs preferentially at the most electron-rich double bond (highlighted in yellow). (II) In AD, the most sterically accessible olefin site (yellow) is dihydroxylated preferentially. (III) Olefinic alcohols are epoxidized site-selectively with $Mo(CO)_6$, $VO(acac)_2$, or $Ti(DET)(O-i-Pr)_2$. acac, acetylacetonate; DET, diethyl tartrate.

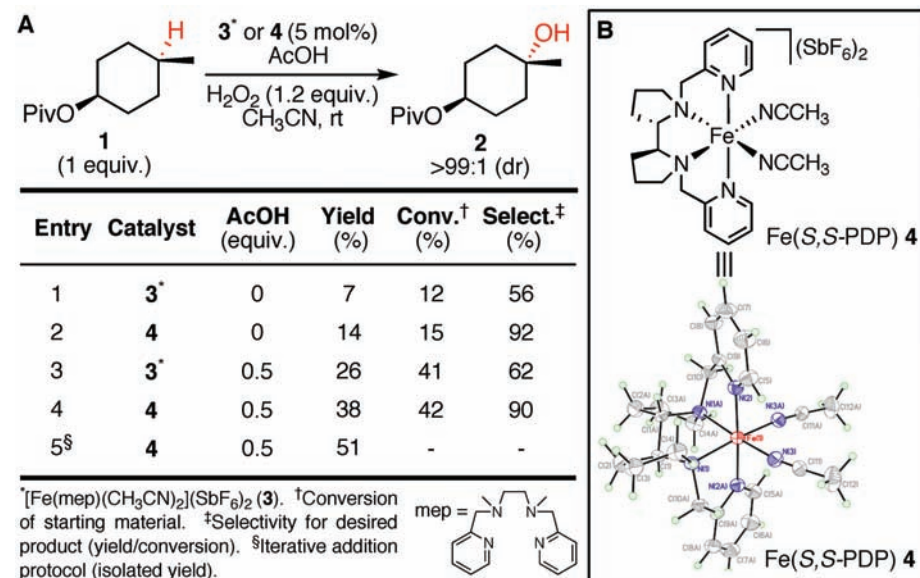
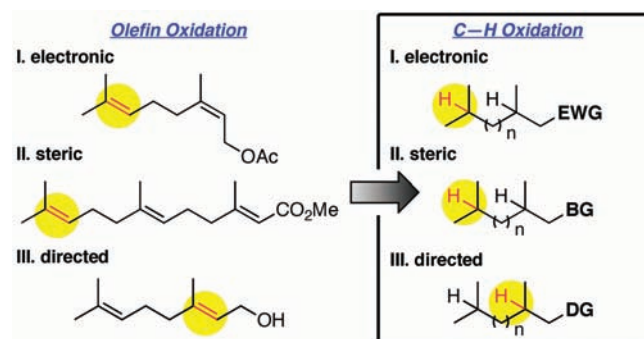


Fig. 2. (A) Development of a preparatively useful aliphatic C–H oxidation reaction. Products resulting from unselective and overoxidation were observed in trace amounts by 1H nuclear magnetic resonance (NMR) and gas chromatograph analysis of the crude reaction mixture. (B) Structure of $[Fe(S,S-PDP)(CH_3CN)_2](SbF_6)_2$ catalyst (**4**) based on x-ray crystallographic analysis (anions are omitted for clarity). PDP indicates 2-((S)-2-((S)-1-(pyridin-2-ylmethyl)pyrrolidin-2-yl)pyrrolidin-1-yl)methylpyridine.

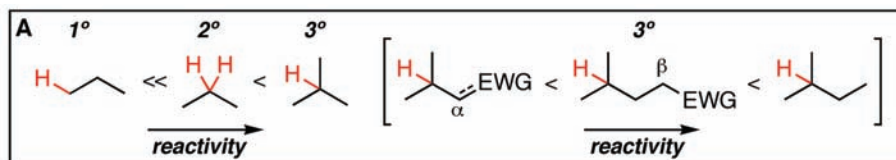
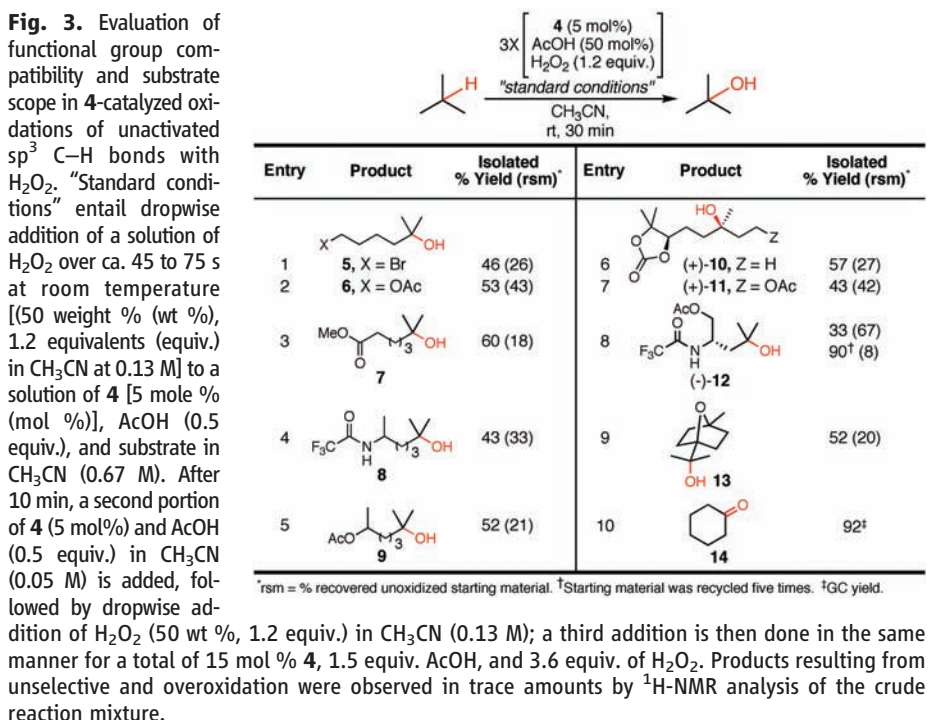
When coupled to asymmetric alkylation methods for constructing stereogenic 3° alkyl centers, this reaction enables a very simple approach for

accessing optically pure tertiary alcohols. In substrates in which no 3° C–H bonds were available, oxidation occurred at the methylene

hydrogens to afford ketone product via the intermediacy of a 2° alcohol (entry 10). The site selectivities and stereochemical outcome of oxidations with **4** are consistent with a concerted mechanism mediated by an electrophilic oxidant (27).

The mass balance of these reactions [average of circa (ca.) 51% mono-oxygenated product with ca. 29% recovered starting material (rsm)] indicates that substantial levels of indiscriminate oxidation are not incurred (Fig. 3). With a highly oxidized *L*-leucinol derivative, hydroxylation occurred exclusively at the 3° C–H bond (entry 8). Although greater than three iterations of **4**, H₂O₂, and AcOH fail to increase product yields, recycling of isolated starting material provides an effective strategy for obtaining high yields with valuable substrates. For example, the *L*-leucinol derivative was recycled five times to obtain a 90% isolated yield of pure (–)-**12** (entry 8).

Complex small molecules often contain multiple 3° C–H centers. We sought to investigate whether the site selectivity of oxidation with electrophilic catalyst **4** is sensitive to the electronic environment of the 3° C–H bond (Fig. 4A). A series of dihydrocitronellol derivatives were evaluated with electron withdrawing groups (EWGs) in α or β positions to one of the two 3° C–H centers (Fig. 4B). In substrates



B

Entry	Substrate	Major Product	Isolated %Yield* (rsm)†	[Remote: Proximal]‡
1	remote proximal	15 , X = H	48§ (29)	1:1
2		16 , X = OAc	43 (35)	5:1
3		17 , X = Br	39 (32)	9:1
4		18 , X = F	43 (20)	6:1
5		19 , X = OAc	49 (21)	29:1
6		20 , X = Br	48 (17)	20:1
7		21 , R = CH ₃	52 (18)	>99:1
8		22 , R = OCH ₃	56 (32)	>99:1

*Unless otherwise noted, isolated yields are of pure major product isolated from the entire reaction mixture. †rsm = % recovered unoxidized starting material. ‡GC analysis of crude reaction mixture using authentic standards. §Isolated as a 1:1 mixture of remote:proximal.

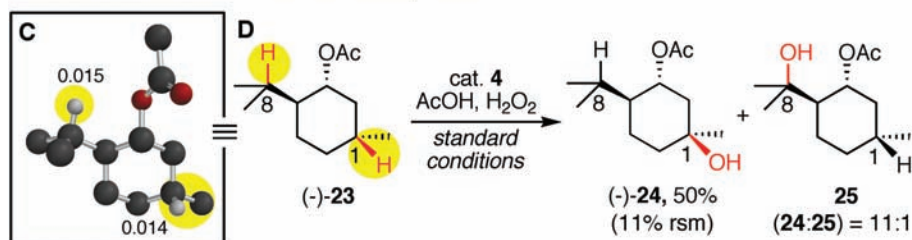


Fig. 4. (A) Reactivity trends for oxidations catalyzed by **4** based on the electronics of the C–H bond. (B) Substrate electronic effects on site selectivity in hydroxylations of multiple 3° C–H bonds with **4**. Only small amounts of diol byproducts were observed. (C) DFT-calculated three-dimensional structure of the lowest potential energy conformer of (–)-**23** with corresponding calculated electrostatic atomic charges (eV) of the 3° C–H bonds of interest. (D) Selective hydroxylation of (–)-**23** at C-1 with **4** based on steric effects. For standard conditions see, Fig. 3. Aliphatic C–H bonds that are oxidized to form product are indicated in red.

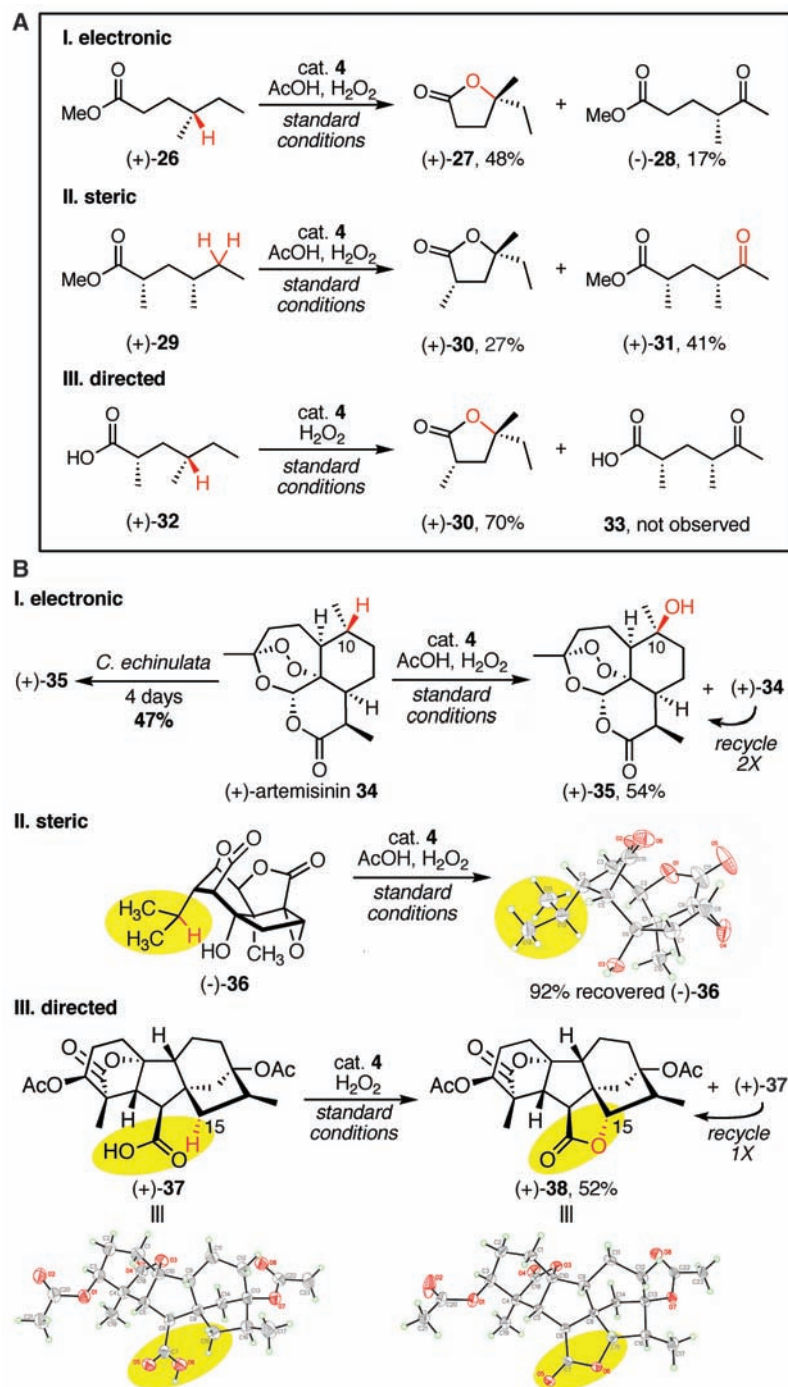


Fig. 5. (A) Three modes of selective aliphatic C–H bond oxidation catalyzed by **4**. Aliphatic C–H bonds that are oxidized to form major product are indicated in red. (I) Oxidation occurs preferentially at the most electron-rich 3° C–H bond followed by in situ lactonization. Unoxidized (+)-**26** was recovered in 23% yield from the reaction. (II) Oxidation occurs at the least sterically hindered, most electron-rich methylene site. Unoxidized (+)-**29** was recovered in 16% yield from the reaction. (III) Oxidation is directed to the sterically hindered 3° C–H site by the free carboxylic acid. (B) Predictably selective aliphatic C–H bond oxidations with **4** of natural products and their derivatives. (I) Selective oxidation of **34** with small molecule catalyst **4** and with cultures of *C. echinulata* occurs at the most electron-rich and least sterically hindered 3° C–H bond to furnish (+)-**35**. (II) Structure of (–)-**36**, determined by x-ray analysis. When (–)-**36** was exposed to standard reaction conditions, 92% of the starting material was recovered because of electronic deactivation of the core and steric deactivation of the isopropyl 3° C–H bond. (III) Carboxylate-directed lactonization of tetrahydrogibberellic acid analog (+)-**37** via C–H oxidation to form lactone (+)-**38** in 52% isolated yield (recycled once). The structures of (+)-**37** and (+)-**38** were determined by x-ray crystallographic analysis and are shown below. For substrates with carboxylic acid directing groups [i.e., (+)-**32** and (+)-**37**], AcOH additive was omitted. For acid-sensitive substrates [i.e., (+)-**34**], AcOH additive was lowered to 10 mol % per addition.

with no electronic bias, equimolar mixtures of hydroxylated products at both centers were formed (entry 1). In all other cases evaluated, hydroxylation with **4** and H₂O₂ occurred preferentially at the 3° C–H bond remote from the EWGs (entries 2 to 8). β-Acetate or halogen functionalities gave modest but useful site selectivities (entries 2 to 4), and α-electron withdrawing functionalities resulted in excellent selectivities for remote hydroxylation (entries 5 and 6). Site selectivities of >99:1 were observed when strongly electron-withdrawing carbonyls were incorporated in the α position relative to one of the 3° C–H bonds (entries 7 and 8). These results demonstrate that C–H oxidations with **4** are subject to electronic deactivation with EWGs in the α or β positions.

We next investigated whether site selectivities of oxidation with bulky catalyst **4** are sensitive to the steric environment of the 3° C–H bond. We chose to examine (–)-acetoxyp-menthane, **23** (Fig. 4D). Energy minimization calculations were performed on (–)-**23** with use of density functional theory (DFT) followed by calculation of the electrostatic atomic partial charges. In the lowest potential energy conformer of (–)-**23**, the two 3° C–H bonds in the γ position to the acetate group (C-1 and C-8) are the least positive and have very similar atomic charges, suggesting a high similarity in their electron densities (Fig. 4C). Thus, only on the basis of the electronic factors, equivalent levels of oxidation would be predicted at these sites. However, we observed a strong preference for oxidation at the C-1 site, most likely because this site is less sterically hindered (Fig. 4D). The gem-dimethyl group of the isopropyl unit in the energy-minimized structure is oriented away from the acetate moiety to relieve unfavorable steric interaction. This conformation places the 3° C–H bond of C-8 proximal to the acetate group, making it sterically less accessible to the oxidant than the C-1 bond (Fig. 4C). These results demonstrate that, in molecules where C–H bonds of similar electron densities are present, sterics can provide a second handle for selectivity.

The interplay between electronic and steric factors in determining the selectivities of C–H oxidations with **4** was further illustrated in a study of methyl esters (Fig. 5A). Hexanoate (+)-**26** was hydroxylated by **4** and H₂O₂ predominantly at the 3° C–H site to afford, after an in situ lactonization, (+)-**27** as the major product with methyl ketone (–)-**28** as the minor product (Fig. 5A). This outcome was predicted on the basis of electronic effects. Increasing the steric bulk around this site by introducing a second methyl substituent in substrate (+)-**29** reverses the selectivity and results in formation of methyl ketone (+)-**31** as the major product. The 2° C–H bond oxidation by **4** also occurred at the most electron-rich, least sterically hindered site. This experiment shows that steric effects can override electronic effects in site

selectivities of oxidation with **4** and suggests that oxidation at 2° C–H sites may operate with selectivities similar to those outlined above for 3° C–H sites.

On the basis of the known role of carboxylates as ligands for nonheme iron complexes and the beneficial role of acetic acid on the catalytic activity of **4**, we postulated that a carboxylate group on the substrate could be used to direct the site of C–H oxidation (28). In support of this hypothesis, hexanoic acid (+)-**32** furnished only the five-membered ring lactone (+)-**30** in 70% isolated yield, whereas oxidation of the analogous methyl ester (+)-**29** gave methyl ketone (+)-**31** as the major product (Fig. 5A). The terminal carboxylic acid moiety in (+)-**32** overrides the previously noted steric effects and directs hydroxylation to the hindered 3° site. Although a unique aspect of aliphatic C–H oxidations catalyzed by **4** is that they do not require a directing group for high selectivities, the ability to use this effect provides a third and powerful handle for selectivity. Predictable reactivity in response to all three such modes of selectivity has proven elusive in prior metal-catalyzed aliphatic C–H oxidations.

The value of this aliphatic C–H oxidation reaction for late-stage synthesis rests on how predictive the electronic, steric, and carboxylate-directing modes of selectivity are in complex molecular settings. In order to evaluate this question, we examined the C–H oxidation of several natural products and their derivatives with **4**. Antimalarial compound (+)-artemisinin **34** displays five 3° C–H bonds along its tetracyclic skeleton (Fig. 5B). In addition to the site-selectivity issue posed in this substrate, a chemo-selectivity challenge is present in the form of a sensitive endoperoxide moiety known to be prone to Fe(II)-mediated cleavage (29). On the basis of the selectivity rules outlined above, we predicted that the electron-rich and sterically unencumbered 3° C–H bond at C-10 would be oxidized preferentially. The remaining 3° C–H bonds are in an α and/or β position to electron-withdrawing ester and endoperoxide moieties. We were gratified to find that the selectivity rules for oxidations with **4** developed on relatively simple substrates could be extended to this complex natural product. (+)-10 β -Hydroxyartemisinin, **35**, was generated as the major product in 34% yield (41% rsm, Fig. 5B). By recycling this valuable starting material through the reaction twice, we obtained diastereomerically pure (+)-**35** in 54% isolated yield. With the same protocol, catalyst **3** afforded (+)-**35** in only 23% isolated yield. Interestingly, (+)-**34** has previously been enzymatically transformed to (+)-**35** in 47% yield with microbial cultures of *Cunninghamella echinulata* (29). Catalyst **4** gives higher yields than the enzymatic reaction with substantially shorter reaction times (three 30-min reactions versus 4 days) and a 10-fold higher volume throughput (0.033 M versus 0.0035 M). The ability of a simple, small mole-

cule catalyst with broad substrate scope to achieve P-450-like tailoring enzyme selectivities is remarkable. We expect that **4** will find widespread use for oxidative modifications to the core structures of natural products and pharmaceuticals.

The high sensitivity of catalyst **4** to steric effects is further illustrated in the attempted oxidation of (–)- α -dihydroprocotoxinin, **36** (Fig. 5B). All of the C–H bonds on the highly oxygenated core are electronically deactivated toward oxidation. Thus on the basis of electronic factors alone, the 3° C–H bond of the exocyclic isopropyl moiety should undergo selective oxidation with 4/H₂O₂. However, treating (–)-**36** under standard hydroxylation conditions resulted in 92% recovered starting material. Examination of an x-ray structure reveals that the isopropyl moiety, in order to avoid severe unfavorable steric interactions, is oriented with its *gem*-dimethyl group projecting away from the ring system (Fig. 5B). This conformation orients the isopropyl 3° C–H bond underneath the ring and renders it inaccessible to **4**. The demonstrated stability of a densely functionalized natural product derivative to this oxidation reaction serves to underscore the remarkably mild nature of this method.

A powerful application of **4**-catalyzed hydroxylations is to effect carboxylate-directed, diastereoselective lactonizations at 2° C–H sites. We chose to evaluate this application with tetrahydrogibberellic acid analog, (+)-**37** (Fig. 5B). The carboxylate moiety of (+)-**37** may direct five-membered ring lactonizations by **4** to one of four C–H bonds. On the basis of x-ray crystallographic analysis of (+)-**37**, we predicted that hydroxylation would occur selectively at the 15 α 2° C–H bond on the D-ring that is closest to the carboxylate moiety. We were gratified to find that hydroxylation with **4** and H₂O₂ furnished five-membered ring lactone (+)-**38** as a single diastereomer in 52% isolated yield (recycled once). With the same protocol, catalyst **3** afforded (+)-**38** in only 26% isolated yield. This large difference in yields with catalyst **4** versus **3** has proven general with all complex substrates examined. Importantly, oxidation of the corresponding methyl ester of **37** resulted in mostly recovered starting material and mixtures of undefined oxidation products, none of which is (+)-**38**.

Given the predictable reactivity and broad scope demonstrated in this study, we anticipate that this general aliphatic C–H oxidation reaction, and others like it, will fundamentally alter the ways in which complex molecules and pharmaceuticals are synthesized in the laboratory.

References and Notes

- K. C. Nicolaou, E. J. Sorensen, *Classics in Total Synthesis* (VCH, New York, 1996).
- K. J. Fraunhoffer, D. A. Bachovchin, M. C. White, *Org. Lett.* **7**, 223 (2005).
- R. W. Hoffmann, *Synthesis* **2006**, 3531 (2006).
- A. R. Dick, M. S. Sanford, *Tetrahedron* **62**, 2439 (2006), and references therein.

- P. Muller, C. Fruit, *Chem. Rev.* **103**, 2905 (2003), and references therein.
- H. M. L. Davies, R. E. J. Beckwith, *Chem. Rev.* **103**, 2861 (2003), and references therein.
- D. J. Covell, N. A. Vermeulen, N. A. Labenz, M. C. White, *Angew. Chem. Int. Ed.* **45**, 8217 (2006).
- K. J. Fraunhoffer, M. C. White, *J. Am. Chem. Soc.* **129**, 7274 (2007), and references therein.
- P. A. Wender, M. K. Hilinski, A. V. W. Mayweg, *Org. Lett.* **7**, 79 (2005).
- S. Lee, P. L. Fuchs, *J. Am. Chem. Soc.* **124**, 13978 (2002). An etheral C–H oxidation with a complex steroidal substrate is shown.
- H. M. L. Davies, X. Dai, M. S. Long, *J. Am. Chem. Soc.* **128**, 2485 (2006).
- A. Hinman, J. Du Bois, *J. Am. Chem. Soc.* **125**, 11510 (2003).
- R. Breslow *et al.*, *J. Am. Chem. Soc.* **95**, 3251 (1973).
- S. Das, C. D. Incarvito, R. H. Crabtree, G. W. Brudvig, *Science* **312**, 1941 (2006).
- J. A. Labinger, J. E. Bercaw, *Nature* **417**, 507 (2002), and references therein.
- B. H. Brodsky, J. Du Bois, *J. Am. Chem. Soc.* **127**, 15391 (2005).
- R. W. Murray, R. Jayaraman, L. Mohan, *J. Am. Chem. Soc.* **108**, 2470 (1986).
- H. C. Kolb, M. S. VanNieuwenhze, K. B. Sharpless, *Chem. Rev.* **94**, 2483 (1994).
- K. B. Sharpless, *Tetrahedron* **50**, 4235 (1994).
- Z. Hu, S. M. Gorun, in *Biomimetic Oxidations Catalyzed by Transition Metal Complexes*, B. Meunier, Ed. (Imperial College Press, London, 2000), vol. 269.
- For the first example of a nonheme iron catalyst demonstrated to effect stereospecific alkane hydroxylation, see C. Kim, K. Chen, J. Kim, L. Que Jr., *J. Am. Chem. Soc.* **119**, 5964 (1997).
- For the first report of an iron(mep) complex and its oxidation reactivity, see T. Okuno, S. Ito, S. Ohba, Y. Nishida, *J. Chem. Soc. Dalton Trans.* **1997**, 3547 (1997).
- K. Chen, L. Que Jr., *Chem. Commun.* **1999**, 1375 (1999).
- M. C. White, A. G. Doyle, E. N. Jacobsen, *J. Am. Chem. Soc.* **123**, 7194 (2001).
- J. England, G. J. P. Britovsek, N. Rabadi, A. J. P. White, *Inorg. Chem.* **46**, 3752 (2007).
- Materials and methods are available at Science Online.
- R. D. Bach, J. L. Andres, M.-D. Su, J. J. W. McDouall, *J. Am. Chem. Soc.* **115**, 5768 (1993).
- E. Y. Tshuva, S. J. Lippard, *Chem. Rev.* **104**, 987 (2004).
- J. Zhan, H. Guo, J. Dai, Y. Zhang, D. Guo, *Tetrahedron Lett.* **43**, 4519 (2002).
- We gratefully acknowledge J. Baudry for DFT calculations, S. Wilson for x-ray crystallographic analysis, M. A. Bigi and D. J. Covell for checking our experimental procedure, and S. E. Denmark and C. S. Regens for a gift of (S,S)-2,2'-bipyrrroline tartrate. M.S.C. is a Harvard University graduate student completing doctoral work with M.C.W. at University of Illinois at Urbana-Champaign. We are grateful to the A.P. Sloan Foundation, the Camille and Henry Dreyfus Foundation, and the University of Illinois for financial support. M.C.W. is a 2007 Lilly Grantee Award recipient. Supplementary crystallographic data for this paper can be obtained from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif on quoting registry nos. CCDC-661933 for **4**, CCDC-661935 for (–)-**36**, CCDC-661934 for (+)-**37**, and CCDC-661936 for (+)-**38**.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/783/DC1

Materials and Methods

Fig. S1

Tables S1 to S4

References

31 July 2007; accepted 1 October 2007
10.1126/science.1148597

4D Visualization of Transitional Structures in Phase Transformations by Electron Diffraction

Peter Baum, Ding-Shyue Yang, Ahmed H. Zewail*

Complex systems in condensed phases involve a multidimensional energy landscape, and knowledge of transitional structures and separation of time scales for atomic movements is critical to understanding their dynamical behavior. Here, we report, using four-dimensional (4D) femtosecond electron diffraction, the visualization of transitional structures from the initial monoclinic to the final tetragonal phase in crystalline vanadium dioxide; the change was initiated by a near-infrared excitation. By revealing the spatiotemporal behavior from all observed Bragg diffractions in 3D, the femtosecond primary vanadium–vanadium bond dilation, the displacements of atoms in picoseconds, and the sound wave shear motion on hundreds of picoseconds were resolved, elucidating the nature of the structural pathways and the nonconcerted mechanism of the transformation.

When transformations of matter involve many atoms, as in complex molecular structures or in condensed phases, understanding their dynamical behavior requires the determination of actual transitional structures with spatial and temporal resolution of the atoms

and their motions. In general, such transformations involve an energy landscape described by transition states and transient intermediates (1), and only when the time scale of observation is appropriate can such species be studied (2). For direct probing, the radiation or particle used must

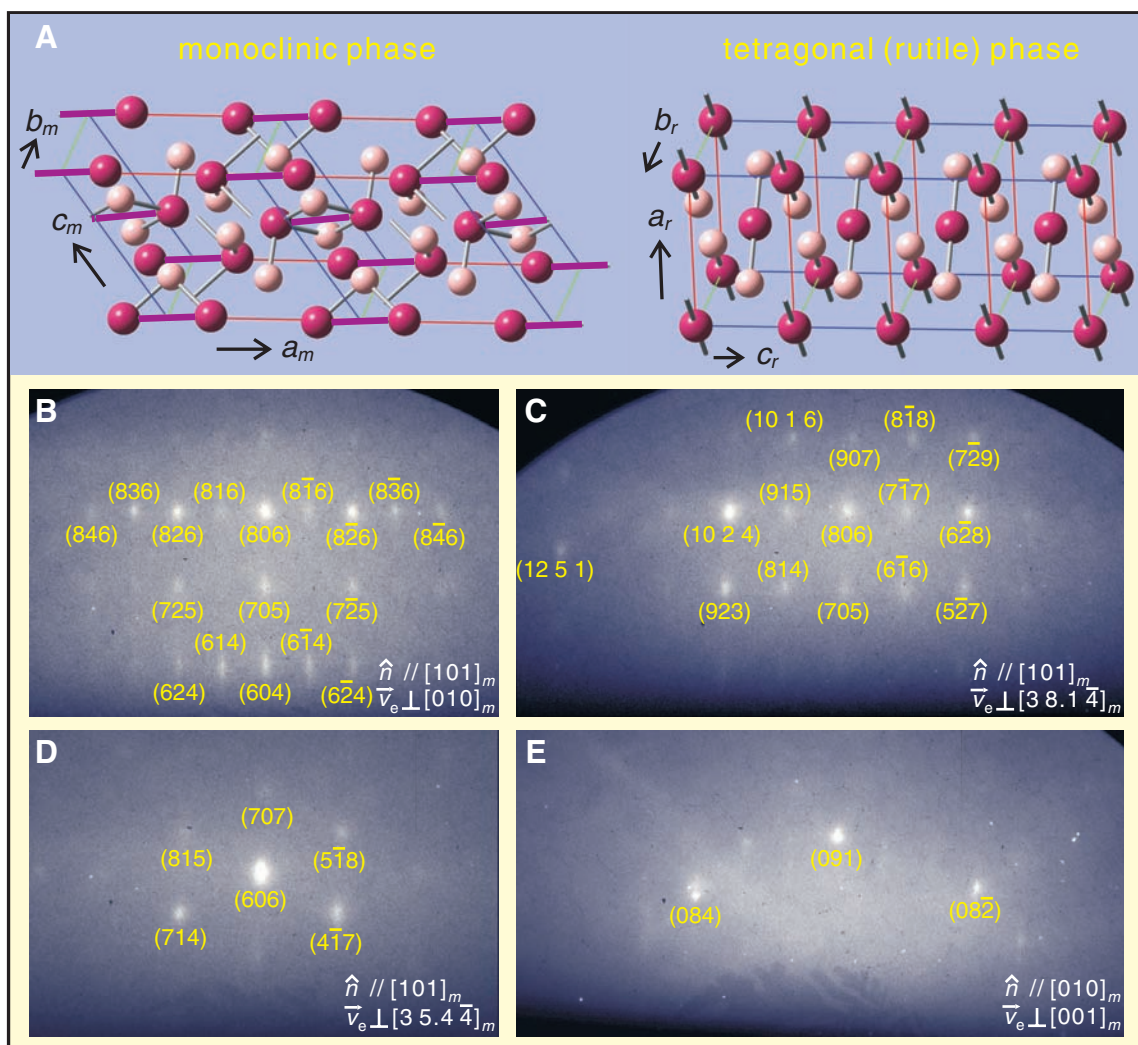
also have a wavelength on the scale of atomic distances, as demonstrated in the studies of melting and lattice dynamics by ultrafast x-ray absorption (3) and diffraction (4–7) and electron diffraction (8–10). In our laboratory, ultrafast electron microscopy (UEM) and ultrafast diffraction (11) have been the methods of choice for studies of molecular and phase transitions (12–16).

In the condensed phase, the use of the energy landscape concept has recently been theoretically addressed (17, 18) for transformations involving two thermodynamically stable configurations, as in solid-solid transitions. There are two classes of descriptions: those that invoke first the motion of the atoms in the unit cell and then the lattice organization and those that deal with the total rearrangement of the lattice including the displacement of unit-cell atoms. The time scales and nature of structures dictate the validity of the theoretical approach and what approximations

Physical Biology Center for Ultrafast Science and Technology, Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed. E-mail: zewail@caltech.edu

Fig. 1. Crystal structures of vanadium dioxide phases and observed Bragg diffraction. (A) Crystal structure of the monoclinic, low-temperature phase (left) and the rutile, high-temperature phase (right). Vanadium atoms are depicted in red and oxygen atoms in a lighter color. Because of a symmetry-breaking pairing of vanadium atoms (violet lines), the axis definitions change. Monoclinic coordinates and axes are used in the paper. (B to E) Typical diffraction patterns observed by ultrafast electron crystallography for different crystal surfaces and different zone axes; \hat{n} is the surface normal direction and \vec{v}_e is the electron direction (zone axis). All Bragg spots can be identified as the monoclinic structure (yellow labels).



are appropriate for the separation of different types of nuclear motions. As with chemical reactions, the concept of concerted (or concurrent) versus consecutive nuclear motions (*1, 2*) becomes central to understanding the elementary steps of the mechanism.

Here, we report on the nature of transitional structures during a symmetry-raising process, from initial monoclinic (insulator) to final tetragonal (metal), observed with atomic-scale spatial and temporal resolution. The study is performed on single crystals of vanadium dioxide, whose phase transition exhibits a well-defined hysteresis between two thermodynamically stable structures. In order to map pathways of motion, all observed Bragg diffractions of different planes and zone axes were examined on the femtosecond to nanosecond time scale. The three-dimensional (3D) sampling and long-range order studied make possible the separation of different nuclear motions, which are mirrored in the temporal change of the structure factor for various indices. Because the transformation takes place in a strongly correlated system, the dependence on excitation fluence is evident in a threshold behavior, and we studied such dependence at short and long times to elucidate the nonequilibrium transition from local atomic motions to shear at sound wave (and carrier) velocity.

At equilibrium, the phase transition in vanadium dioxide has been studied by examining the change of heat capacity with temperature (*19–21*). Such studies identify the transition as first order with a hysteresis. Seminal contributions using optical reflection and x-ray methods have indicated the ultrafast nature of the

transition (*22–26*). The transition can be induced nonthermally on the ultrashort time scale, with the required excitation a variable characteristic dependent on the morphology of the specific sample. Real-space imaging (and diffraction) was achieved with UEM for nanoscale structures (*14, 15*). We investigated high-quality single crystals of vanadium oxide, invoked ultrafast electron crystallography with a tilted geometry (*11, 27, 28*), and used the 3D reciprocal space of atoms involved (20 Bragg spots).

The initial and final crystal structures of vanadium dioxide are depicted in Fig. 1A: in the monoclinic phase the vanadium atoms arrange into pairs, but in the tetragonal phase (referred to as rutile) all V–V distances are equal and the symmetry is tetragonal (*21*). The structure and long-range order of our vanadium dioxide single crystals were confirmed by static x-ray diffraction, and one of three crystals was cut and polished to have a surface that is not naturally grown. For these crystals, the characteristic hysteresis near 340 K was observed, with a width of 5.6 K. In Fig. 1, B to E, the static electron diffraction patterns, obtained at an incidence angle of $\sim 5^\circ$ and at room temperature (monoclinic phase), are shown for different surface normal directions, \hat{n} , and electron directions (zone axis), \vec{v}_e . All spots were identified as monoclinic vanadium dioxide; transmission-like patterns originate from surface structures (*12*), and the single-crystal order is evident in the well-indexed diffraction patterns.

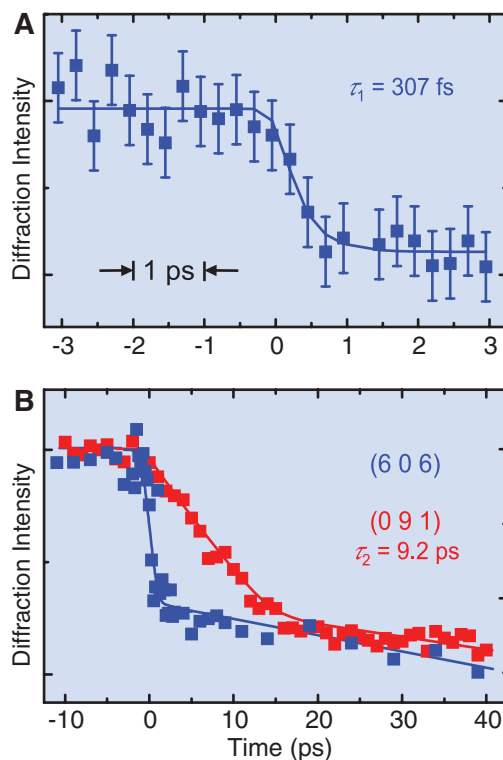
With such rich diffraction, 16 of the Bragg spots were intense enough for time-resolved investigations. The change was initiated by using

near-infrared (800 nm) pulses, and, after a variable delay time, structural dynamics were followed by diffracting the electron packets. The optical excitation fluence was up to 14 mJ/cm^2 , enough to drive the phase transition with a single pulse. Within the experimental repetition period of 1 ms, the crystal held at room temperature fully recovers to the initial monoclinic phase. This recovery was confirmed by observing no change in the diffraction patterns whether recorded at negative delay time (an effective 1-ms delay) or without the excitation. The observed changes by fs excitation, therefore, reveal the nonequilibrium dynamics without contributions from static heating.

We checked for possible effects of surface potential change or charge trapping on the diffraction during the transition from insulator to metal phase; no such effects were found, as evidenced from our observation of a steady position and intensity of the direct nondiffracted electron beam for scans at all time delays. At 5° incidence, the electrons diffract from a material thickness of about 10 nm; we do not observe rods in the diffraction patterns but instead well-defined Bragg spots, suggesting that at least 10 interatomic layers in the surface-normal direction are contributing to the interferences (*28*). A rough estimate of the mean free path of a 30-keV electron gives a penetration depth ($\sim 5 \text{ nm}$) that is about 20 times larger than the interplanar separation. Because of the small angle of incidence, the probed area on the surface is $\sim 2 \text{ mm}$ by 0.2 mm , giving a group velocity mismatch of $\sim 20 \text{ ps}$ between the optical excitation and the electron pulses. In order to overcome such mismatch and reach fs resolution, the optical pulse was tilted to achieve temporal synchrony with the electron packet on the entire probed crystal surface (*27*). A fs transient, however, includes convolution from any residual spread and the involved duration of optical and electron pulses, as discussed below.

The observed temporal change of intensity of one class of Bragg spots, for example, the (606) spot, is shown in Fig. 2A. With time steps of 250 fs, a notable decrease in intensity is observed. In order to record such a transient, the number of electrons in one pulse was reduced to ~ 500 , which is below the space-charge limit (*11, 13*), and in this limit of low flux the electron pulse width was measured ($322 \pm 128 \text{ fs}$) in situ at a streaking speed of $140 \pm 2 \text{ fs per pixel}$ (*29*). Given the optical pulse width of 120 fs, the cross-correlation was used in the analysis of the transient, thus obtaining a time constant, τ_1 , of 307 fs. The error bar at each delay time is a result of 18 single measurements. For all of the Bragg spots studied (Fig. 1), we observed two types of behavior (Fig. 2B): a fs decay in intensity (blue) similar to the (606) spot and, remarkably, a decay lacking such fast dynamics but exhibiting an intensity decrease with a time constant, τ_2 , of 9.2 ps (red); the Miller indices (*hkl*) of the two classes of spots are listed in the figure caption. We note that no shift in position or change in width

Fig. 2. Ultrafast, fs and ps, diffraction dynamics of the structural phase transition. **(A)** Intensity change of the (606) Bragg spot with time. A decay with a time constant τ_1 of 307 fs was obtained on the total time scale of $\pm 3 \text{ ps}$ (*40*); note the change in intensity of diffraction in the 250-fs steps. **(B)** Intensity change of (606) (blue) and (091) (red) spots with time. For all investigated Bragg spots, two different types of dynamics were observed: a femtosecond decay similar to the blue trace was measured for (806), (826), ($8\bar{2}6$), (846), ($8\bar{4}6$), (606), (714), ($4\bar{1}7$), ($10\bar{2}4$), and ($6\bar{2}8$); a decay with a time constant τ_2 of 9.2 ps, similar to the red trace, was measured for (091), (084), and ($08\bar{2}$), on the total time scale of 40 ps. The temporal range values (Δt_1 and Δt_2) over which the decay is pronounced are 760 fs and 15 ps (*40*). This difference indicates a stepwise mechanism for atomic motions.



was measurable at early times. The absence of a spot shift (12) shows that no substantial lattice expansion is taking place on the ultrashort time scale; the absence of a clear width change also indicates that no measurable disorder is introduced on this time scale. These observations exclude thermal expansion or lattice inhomogeneity (strain) in this time range (28).

All Bragg spots that show the fs behavior involve nonzero values of (hkl) , whereas those displaying the slower ps behavior have a zero component of h . Accordingly, the observed fs and ps intensity changes of Bragg spots are associated with motion of atoms within the unit cell by destructive interferences, and these changes are determined by the structure factor monitored in the intensity of scattered electrons for a given direction and zone axis. Specifically, the intensity I of a Bragg spot (hkl) , which is proportional to the square of the structure factor $F(hkl)$, is determined by the position (xyz) of atom j within the unit cell (Eq. 1)

$$F(hkl) = \sum_j f_j \exp[-2\pi i (hkl) \cdot (xyz)_j]$$

$$I(hkl) \propto |F(hkl)|^2 \quad (1)$$

where f_j is the atomic scattering factor. The monoclinic phase has a lower symmetry than that of the rutile phase, and the symmetry-raising processes cause a large change in Bragg diffraction, even when small-amplitude atomic motions are involved, as shown below.

The observed two types of dynamics indicate stepwise atomic motions along different directions. Figure 3A depicts the initial and final vanadium positions in the two phases. From the inner product in Eq. 1, it is evident that an atomic movement along a certain direction can only affect such Bragg spots that have nonzero contributions in the corresponding Miller indices. It is thus concluded that the initial fs motion is along the a axis, which is the direction of the V–V bond in the monoclinic structure. On the other hand, the ps structural transformation projects along the c and b axes (Fig. 3A). If the fs motion had large components along b or c , it would show up in the dynamics of all investigated spots, contrary to observations.

To quantify selective intensity changes, we calculated for the observed Bragg diffraction the structure factor and intensity changes of the monoclinic phase when the atoms undergo small atomic displacements. Two possible pathways for atomic movement were considered. In Fig. 3A, we display the consequence of V–V bond weakening (or dilation) motion on the values of $|F|^2$; all spots with $h \neq 0$ show a decrease in intensity, and for those with $h = 0$ the intensity remains unchanged, as observed experimentally. In contrast, when considering direct movement leading to the transition from the monoclinic to the rutile phase, most spots show an increase and some a decrease in $|F|^2$ values, a behavior that was not observed experimentally on the ultrafast time scale.

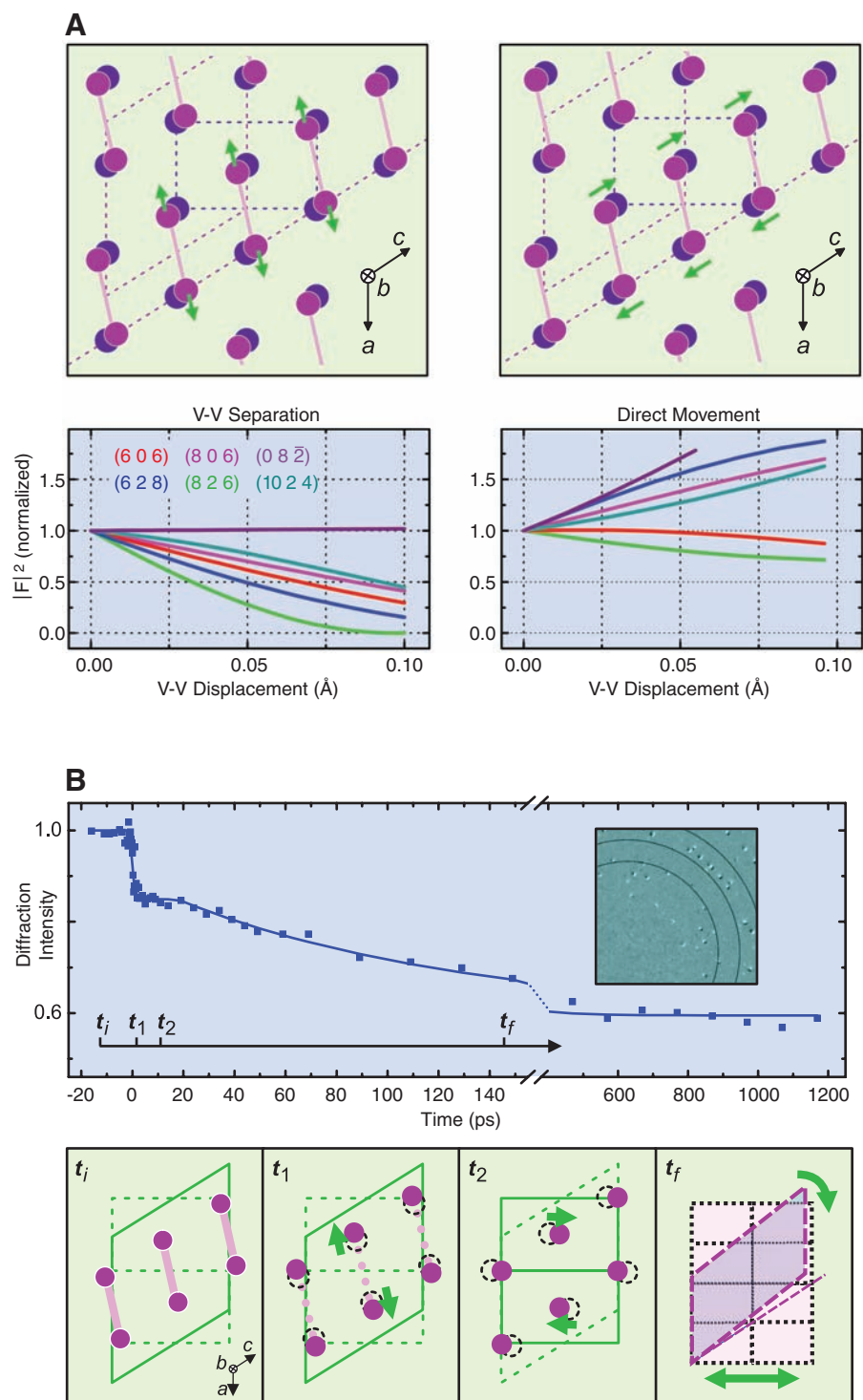


Fig. 3. (A) Atomic movements and the calculated structure factors. (Top left) Initial (purple) and final (blue) positions of the vanadium atoms. The monoclinic unit cell is depicted as dotted purple lines, and the rutile unit cell as a blue dotted box. (Top right) The direct motion of the vanadium atoms from the initial (purple) to the final (blue) position, mostly along the c axis, is depicted. (Bottom left) Structure factor and expected intensity change for the listed Bragg spots for a displacement of the vanadium atoms along the chemical bond, as depicted in the top left image. (Bottom right) Structure factor and intensity change for a displacement of the vanadium atoms along the direct path, as depicted in the top right image. The V–V displacements shown are for delocalized excitation of the lattice. (B) Transitional structures during the ultrafast phase transformation. (Top) The experimental data, here the (606) spot (blue), reflects the stepwise atomic motion within the unit cell and, on larger length scales, of shear motion. (Inset) The difference in static x-ray patterns between monoclinic (low temperature) and rutile (high temperature) vanadium dioxide suggests macroscopic shear (angular) displacement. (Bottom) Shown are frames at the corresponding times (t_i , t_1 , t_2 , and t_f) of structural changes.

The initiating excitation at 1.55 eV primarily involves the $d_{||}$ band, which arises from bonding of the vanadium pairs (30, 31). From a chemical perspective, the excitation is to an antibonding state, which instantly results in a repulsive force on the atoms, and they separate along the bond direction. In sequence and on a slower time scale, the unit cell transforms toward the configuration of the rutile phase. Therefore, the observed stepwise atomic motions show that the phase transition proceeds by a nondirect pathway on the multidimensional potential energy surface and not by a direct structural conversion.

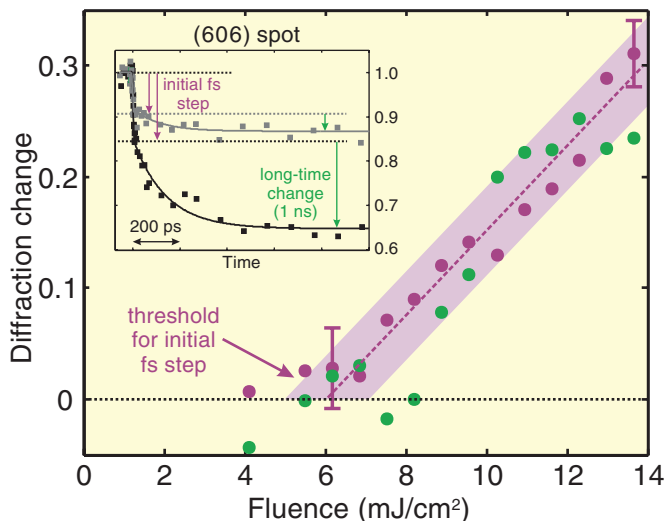
Several points are worth mentioning. First, these resolved structures are for transient species, en route to the stable rutile structure, and involve a landscape with distinct coordinates. In Fig. 3B, the overall temporal behavior up to 1.2 ns and structural snapshots are shown. Second, optical reflectivity and diffraction from polycrystalline samples provide an average over all orientations and will give rise to composite transient behavior. Cavalleri *et al.* (22) have shown that x-ray diffraction around the (011) spot region (32) of a crystal is dominated by ~ 12 -ps change, with a relatively small faster component (≤ 500 fs). Given our results, we concluded that dominance of the fs dynamics can emerge if the relevant direction of diffraction is monitored and that the ps component is due to the transversal motion. Lastly, for the first step, comparison with the ~ 100 -fs time constant from optical reflectivity (23, 25) and ~ 500 fs from the x-ray study (22) is not straightforward, because the former approach is more sensitive to the electronic changes of the material whereas the latter one integrates the structural changes over the x-ray probing length of micrometers. Future experiments may reveal the vibrational time scale of ~ 170 fs of the equilibrium structure or a somewhat longer time due to a potential-driven motion in the excited state (28).

The transient behavior on the longer, subnanosecond time scale reveals another dimension

of structural dynamics. As shown in Fig. 3B, after the initial insulator-metal transition after a delay, there is a continued temporal change of intensity that levels off at ~ 300 ps. Given the optical penetration depth ($1/\alpha$) of ~ 100 nm (20) and the thermal diffusion coefficient of $0.02 \text{ cm}^2/\text{s}$ (20), we conclude that heat conduction away from the probed layer must be slower than 1 ns and cannot explain the observed ~ 100 -ps dynamics. However, because the speed of sound in the material is $v \approx 4000 \text{ m/s}$ (33), shear, which is necessary for the formation of rutile vanadium dioxide (34) and may involve dislocations (35), will occur in ~ 100 ps for a length scale of hundreds of nanometers (36). This shear motion is further supported by other observations. First, although the Bragg spots exhibit the same early fs-ps behavior, they show intensity increase or decrease at the sub-ns scale for different zone axes, consistent with shear interferences (37). Second, when we obtained x-ray diffraction of the crystal at different temperatures (Fig. 3B, inset) as it passed through the phase transition, nearly all Bragg spots were observed to move angularly, indicating rotation of the principle axes. Further evidence comes from the fluence dependence and its threshold behavior.

Figure 4 shows the fluence dependence of diffraction change, for example, for the (606) Bragg spot, at two time points, $t = 10$ ps and $t = 1$ ns. The intensity change displays (almost) a linear dependence with a threshold at $6 \pm 1 \text{ mJ}/\text{cm}^2$. The threshold, which translates to ~ 0.05 photon per vanadium atom (38), indicates the minimum fluence required for switching into the new phase in this crystalline material and as such defines a level of connectivity among the different sites involved (15). For the spots that exhibit the fs dynamics, we noted that the threshold for the sub-ns component is either the same or higher than that of the fs component, suggesting that the shear is associated with the initial atomic motions.

Fig. 4. Dependence of the fs dynamics and the large-scale shear displacement on excitation fluence. The magnitude of the initial intensity decrease (violet) and the amplitude of the long-time change (green) are plotted for different excitation fluences. A common threshold region is observed. (Inset) Transients of the (606) Bragg spot for two different excitation fluences. The magnitudes of change were measured at 10 ps and 1 ns.



From Eq. 1, the calculated V–V displacement (Fig. 3A) reaches 0.02 \AA for ~ 0.1 photon per vanadium for delocalized excitation of the lattice. However, vanadium dioxide is a strongly correlated system, and this value could be as high as 0.2 \AA for localized V–V pairs. We note that the total energy deposited at threshold ($\sim 0.4 \pm 0.1 \text{ J}/\text{mm}^3$) matches well with the total heat required ($0.38 \text{ J}/\text{mm}^3$) for thermally inducing the phase transition, including the latent heat (20). This observation suggests that the phase transition is critically dependent on the total number of carriers, thermal and optical (26), and that the described transitional structures may be reached similarly but with different rates, as shown, for example, in surface femtochemistry (39). Lastly, we have repeated these experiments at different temperatures; at 110 K, the fluence threshold shifts to a higher value compared with that at 300 K, which further supports the above-mentioned stepwise mechanism.

The ability to decipher the nature of the atomic motions during a structural phase transition is demonstrated by the results reported here. For vanadium dioxide crystals, the elementary steps follow a nonconcerted mechanism with a sequence of transitional structures, first involving local displacements on the fs and ps time scale followed by long-range shear rearrangements on the sub-ns time scale and at the speed of sound. The V–V bond dilation is the initial step of the insulator-to-metal transformation, providing a dynamical molecular picture. The coincidence of the thermal and photoinduced transition thresholds at different temperatures suggests the common pathway mechanism for the transition. With 4D atomic-scale spatial and temporal resolutions, we expect, by using this table-top approach, many future extensions in the studies of designed nanoscale materials and biological systems (11).

References and Notes

1. J. C. Polanyi, A. H. Zewail, *Acc. Chem. Res.* **28**, 119 (1995).
2. A. H. Zewail, *Science* **242**, 1645 (1988).
3. C. Bressler, M. Chergui, *Chem. Rev.* **104**, 1781 (2004).
4. A. Rousse, C. Rischel, J.-C. Gauthier, *Rev. Mod. Phys.* **73**, 17 (2001).
5. K. Sokolowski-Tinten, D. von der Linde, *J. Phys. Cond. Matter* **16**, R1517 (2004).
6. M. Bargheer, N. Zhavoronkov, M. Woerner, T. Elsaesser, *ChemPhysChem* **7**, 783 (2006).
7. K. J. Gaffney, H. N. Chapman, *Science* **316**, 1444 (2007), and references therein.
8. H. E. Elsayed-Ali, G. A. Mourou, *Appl. Phys. Lett.* **52**, 103 (1988).
9. B. J. Siwick, J. R. Dwyer, R. E. Jordan, R. J. D. Miller, *Science* **302**, 1382 (2003).
10. H. Park, X. Wang, S. Nie, R. Clinitie, J. Cao, *Phys. Rev. B* **72**, 100301 (2005).
11. A. H. Zewail, *Annu. Rev. Phys. Chem.* **57**, 65 (2006), and references therein.
12. N. Gedik, D.-S. Yang, G. Logvenov, I. Bozovic, A. H. Zewail, *Science* **316**, 425 (2007).
13. V. A. Lobastov, R. Srinivasan, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7069 (2005).

14. M. S. Grinold, V. A. Lobastov, J. Weissenrieder, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18427 (2006).
15. V. A. Lobastov, J. Weissenrieder, J. Tang, A. H. Zewail, *Nano Lett.* **7**, 2552 (2007).
16. H. S. Park, J. S. Baskin, O.-H. Kwon, A. H. Zewail, *Nano Lett.* **7**, 2545 (2007).
17. K. J. Caspersen, E. A. Carter, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6738 (2005).
18. D. R. Trinkle *et al.*, *Phys. Rev. Lett.* **91**, 025701 (2003).
19. F. J. Morin, *Phys. Rev. Lett.* **3**, 34 (1959).
20. C. N. Berglund, H. J. Guggenheim, *Phys. Rev.* **185**, 1022 (1969).
21. K. D. Rogers, *Powder Diff.* **8**, 240 (1993).
22. A. Cavalleri *et al.*, *Phys. Rev. Lett.* **87**, 237401 (2001).
23. G. I. Petrov, V. V. Yakovlev, J. A. Squier, *Opt. Lett.* **27**, 655 (2002).
24. A. Cavalleri *et al.*, *Phys. Rev. Lett.* **95**, 067405 (2005).
25. A. Cavalleri, M. Rini, R. W. Schoenlein, *J. Phys. Soc. Jpn.* **75**, 011004 (2006), and references therein.
26. H.-T. Kim *et al.*, *Phys. Rev. Lett.* **97**, 266401 (2006).
27. P. Baum, A. H. Zewail, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16105 (2006).
28. D.-S. Yang, N. Gedik, A. H. Zewail, *J. Phys. Chem. C* **111**, 4889 (2007).
29. V. A. Lobastov *et al.*, in *Ultrafast Optics IV*, F. Krausz, G. Korn, P. Corkum, I. A. Walmsley, Eds., vol. 95 of *Springer Series in Optical Sciences* (Springer, New York, 2004), pp. 419–435.
30. J. B. Goodenough, *Phys. Rev.* **120**, 67 (1960).
31. T. C. Koethe *et al.*, *Phys. Rev. Lett.* **97**, 116402 (2006).
32. Given the x-ray wavelength (1.54 Å) and diffraction angle (13.9°) reported (22), the Bragg peak (22) should be indexed as (011) of the monoclinic phase. It was assigned as (110) of the monoclinic phase, which in fact becomes (110) of the rutile phase upon transformation.
33. D. Maurer, A. Leue, R. Heichele, V. Müller, *Phys. Rev. B* **60**, 13249 (1999).
34. D. Kucharczyk, T. Niklewski, *J. Appl. Crystallogr.* **12**, 370 (1979).
35. J. M. Thomas, *Chem. Br.* **6**, 60 (1970), and references therein.
36. Excess electrons (carriers), which are not bound by strong correlation with the lattice, may redistribute within the excited metallic region but are impeded by the insulating surrounding. Electron diffusion from the probed layer (~10 nm) to the metallic region (~100 nm) occurs in ~100 ps, as calculated from the electron mobility of 1 to 10 cm²/(V-s) for metallic vanadium dioxide (20). The diffusion of such electrons into the deeper regions may contribute to generation of shear.
37. Shear motion leads to a change in principal axes (34). Because not all Bragg spots are equally well in phase with the Ewald sphere at the same time (28), shear motion may enhance or suppress the Bragg intensities to values above or below the initial intensity, as observed.
38. From the reflectivity of 0.28 and the absorption depth of 100 nm at 800 nm (20), the threshold fluence corresponds to 450 mJ/mm² at the surface. With the unit cell volume of 118 Å³ (21), which contains four vanadium atoms, this energy density gives ~0.05 photon per vanadium.
39. G. Ertl, *Adv. Catal.* **45**, 1 (2000).
40. In order to evaluate the maximum range of the intensity decay, we also considered a convoluted step function instead of a decay process. This distinction becomes significant depending on the physics of the process involved. For a step function, we obtained $\Delta t_1 = 760$ fs. The overall fit of the transient was repeated 1000 times to estimate the error in the Δt_1 range, which was found to be ± 80 fs. We note that changes in intensity occur in a step of 250 fs. For the ps component, the range Δt_2 of 15 ps is evident from the figure.
41. We are grateful to J. Weissenrieder for helpful discussions, L. H. Tjeng for generously providing some crystals, G. R. Rossman for the crystal-cutting equipment, and L. M. Henling for help with the x-ray measurements. This work was supported by the NSF, by the Air Force Office of Scientific Research, and by the Gordon and Betty Moore Center for Physical Biology at Caltech. P.B. was partially supported by the Alexander von Humboldt Foundation.

12 July 2007; accepted 19 September 2007
10.1126/science.1147724

Molecular and Genomic Data Identify the Closest Living Relative of Primates

Jan E. Janečka,¹ Webb Miller,² Thomas H. Pringle,³ Frank Wiens,⁴ Annette Zitzmann,⁵ Kristofer M. Helgen,⁶ Mark S. Springer,⁷ William J. Murphy^{1*}

A full understanding of primate morphological and genomic evolution requires the identification of their closest living relative. In order to resolve the ancestral relationships among primates and their closest relatives, we searched multispecies genome alignments for phylogenetically informative rare genomic changes within the superordinal group Euarchonta, which includes the orders Primates, Dermoptera (colugos), and Scandentia (treeshrews). We also constructed phylogenetic trees from 14 kilobases of nuclear genes for representatives from most major primate lineages, both extant colugos, and multiple treeshrews, including the pentail treeshrew, *Ptilocercus lowii*, the only living member of the family Ptilocercidae. A relaxed molecular clock analysis including *Ptilocercus* suggests that treeshrews arose approximately 63 million years ago. Our data show that colugos are the closest living relatives of primates and indicate that their divergence occurred in the Cretaceous.

The origins of modern primates and their fossil relatives remain a topic of intense debate (1–3), as there has been an increased focus on identifying adaptive evolutionary changes within primates and the dynamics of genome evolution within the primate lineage (4, 5). Resolving higher primate relationships has been challenging, making it difficult to identify character transformations in early primate evolution. An essential part of this challenge is to determine the closest living relative to primates, which would provide a broader context for understanding primate evolution.

DNA sequence and morphological studies, and analyses of rare genomic changes, support

the monophyly of treeshrews, colugos (flying lemurs), and primates in the clade Euarchonta, with a sister-group relationship to Glires [which includes rodents and lagomorphs (3, 6–8)]. In contrast, the relationships within Euarchonta are not well resolved, most likely because of the rapid evolution of these groups and inadequate sampling within Scandentia and Dermoptera. Three hypotheses have been proposed: (i) a sister-group relationship between treeshrews and primates (9–11), (ii) a sister-group relationship between colugos and primates [Primates + (12)], and (iii) both colugos and treeshrews as sister to the primates [Sundatheria (2, 13)]. Molecular and morphological studies have favored

Sundatheria (3, 6, 14), although support for this hypothesis was lower than for other mammalian interordinal clades (15). Primatomorpha, proposed on morphological grounds (12), has also been indicated by some molecular studies (16, 17). Other studies have failed to reject alternative hypotheses, and analyses of different character subsets support contradictory topologies (18–20).

To improve our understanding of early euarchontan evolution and determine the closest living relative of primates, we used two independent molecular approaches. We first screened a nonredundant set of 197,522 protein-coding exons from the human University of California Santa Cruz Known Genes track to identify rare genomic changes (exonic indels) that would potentially support the three a priori hypotheses. We also assembled and analyzed a 14-kb nuclear gene data set of 19 gene fragments in order to estimate a phylogeny and time scale for extant euarchontans. To mitigate against the possible effects of long-branch attraction (LBA), which can incorrectly place rapidly evolving clades

¹Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843, USA. ²Center for Comparative Genomics and Bioinformatics, Pennsylvania State University, University Park, PA 16802, USA. ³Sperling Foundation, Eugene, OR 97405, USA. ⁴Department of Animal Physiology, University of Bayreuth, D-95440 Bayreuth, Germany. ⁵Zoological Institute, Johann Wolfgang Goethe-University, D-60054 Frankfurt/Main, Germany. ⁶National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA. ⁷Department of Biology, University of California, Riverside, CA 92521, USA.

*To whom correspondence should be addressed at the Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Veterinary Medical Administration Building, Room 107, Texas A&M University, College Station, TX 77843–4458, USA. E-mail: wmmurphy@cvm.tamu.edu

together, we included nuclear DNA sequences from both living colugos and the second tree shrew family, Ptilocercidae (21, 22).

We identified 300 candidate indels in coding gene exons within Euarchonta. Of these, 104 were excluded because they lacked flanking sequences that were long or conserved enough for primer design, were determined to be anomalous misalignments, or were computationally determined to be paralogous gene alignments (23). The lack of a colugo genome sequence required polymerase chain reaction (PCR) am-

plification of candidate indel-containing exons in the colugo and comparison to the treeshrew genome sequence (23). PCR primers were designed for the remaining 196 candidates, of which 75% produced a single band in colugo, distributed in the following categories: 32 indels that were initially primate-specific (shared by anthropoids and strepsirrhines, potentially informative for Primatomorpha); 13 indels shared by primates and treeshrews (potentially informative for colugos being in a basal position or alternatively for euarchontan monophyly);

and 102 indels that were treeshrew-specific (potentially informative for Sundatheria).

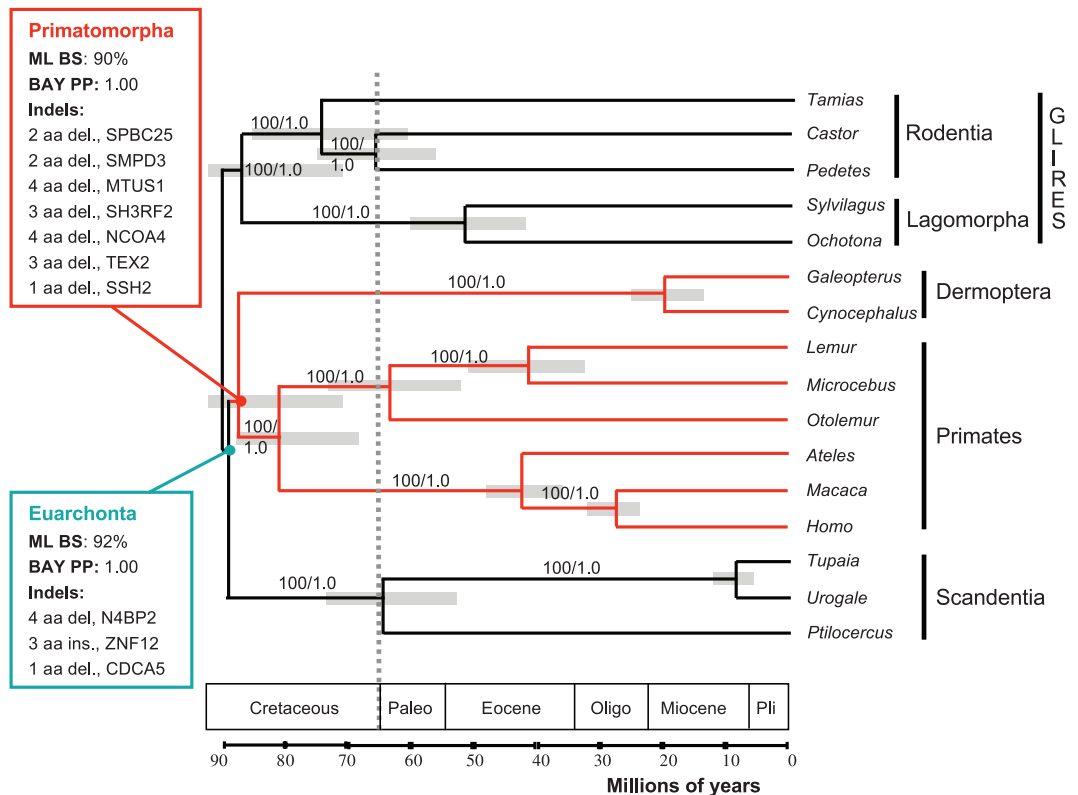
After excluding noninformative and hyper-variable indels (23) and the evaluation of additional eutherian genomes (table S1), three indels supported the monophyly of Euarchonta [*N4BP2*, *ZNF12*, and *CDC45* (figs. S1 to S3)] and corroborate the emerging phylogenetic consensus that primates, colugos, and treeshrews are a monophyletic group (3, 8, 15, 19, 20). No indels placed treeshrews with rodents and lagomorphs (17) or treeshrews as basal within Euarchontoglires (24). We identified seven indels that supported colugos as the closest living relative of primates [Primatomorpha: *SPBC25*, *SMPD3*, *MTUS1*, *SH3RF2*, *NCOA4*, *TEX2*, and *SSH2* (Figs. 1 and 2 and figs. S4 to S10)]. By contrast, no indels supported Sundatheria, despite a larger number of potentially informative candidates for this hypothesis having been screened. One indel (*ADD2*) supported a sister-group relationship between treeshrews and primates (fig. S11). Taken together, an analysis of these last eight indels by means of a statistical framework (7) provides significant support for Primatomorpha [$P < 0.025$ (23)].

The monophyly of Primatomorpha was independently confirmed by phylogenetic reconstruction from a 14-kb data set consisting of 19 nuclear gene segments with maximum likelihood [(ML) 90% bootstrap support] and Bayesian (1.00 posterior probability) algorithms (Fig. 2 and fig. S12). Previously, the hierarchical order at the base of the Euarchonta was difficult to resolve with confidence because of contempo-

Fig. 1. An example of a coding sequence indel supporting the Primatomorpha hypothesis. A three-amino acid deletion in exon 4 of the *TEX2* gene is present in all major primate lineages and both colugo genera (shaded gray) but is absent in all treeshrew lineages and eutherian outgroup representatives. See figs. S1 to S10 for full alignments and descriptions of additional supporting indels for Euarchonta and Primatomorpha.

TEX2 exon 4	
human	SEEKPPAE---GSEDPKKPPRPQEGTR
chimp	SEEKPPAE---GSEDPKKPPRPQEGTR
orangutan	SEEKPPAE---GSEDPKKPPRPQEGTR
macaque	SEEKPPAE---GSEDPKKPPRPQEGTR
marmoset	SEEKPSAE---GSEDPKKPPRPQEGTR
tarsier	SEEKPPPE---GSEDPQKPPPPQEGTR
bushbaby	LEELPAE---GSEDPKKPPHPQEGTR
mouse lemur	LEELPVE---GSEDPKKPPHPQEGAR
Phil. flying lemur	VEELPAE---GSEDPKKPPVPQEGTR
northern treeshrew	SEDKPPAERELGSEDPKKPPHSQEGTR
pentail treeshrew	SEEKPPAEREPPGSEDPKKPPHSQEG-R
mouse	TEEKPPPEKELPSEDLKKPPQPEGTK
guinea pig	SEEKPPAEKELGSEDPKKPSHPQEGTR
rabbit	SEEKPPAERELASEDPKKPPQPEGTR
dog	SEEKPPAERELGGEDPKKPPHPQEGTR
horse	SEEKPPTEKEQGVEDPKKPSPPQEGTR
brown bat	SEEKPPAERDLGVEDPKKPPHPQEGTR
cow	CEEKPPAERELGGEDPKKPPHPQEGTR
shrew	SEEKPPAEKELGAEDPKKPAHPQEGTR
armadillo	SEEKPSAERELGSEDSKKPPHSQEGTR
elephant	SEEKPPAERELAGEDPKKPP--LEGTR

Fig. 2. A maximum-likelihood phylogeny of the superorder Euarchonta, with rodent and lagomorph lineages as outgroups. Branch lengths were estimated under an F84 model of sequence evolution and the relaxed molecular clock approach, implemented in the program MULTIDIVTIME (23). Bootstrap (BS) values and Bayesian posterior probabilities (BPPs) are shown on branches for which these values are 100% and 1.0, respectively. Amino acid (aa) indels (ins, insertion; del, deletion) supporting the monophyly of Euarchonta and Primatomorpha are listed in boxes to the left, along with respective BS and BPP values. A molecular time scale is presented below the tree (23). The 95% credibility intervals (CIs) are shown as gray bars spanning each node. The point estimates and 95% CIs for all nodes are presented in table S4.



aneous divergence of ancestral lineages during the Cretaceous, LBA, and limited taxon and gene sampling (25). The results from our expanded data set (table S2) contrast with previous studies supporting Sundatheria. When *Ptilocercus lowii* and both colugo genera are included, the ML and Bayesian trees become consistent with rare genomic changes. The importance of *P. lowii* was evident when it was removed from the data set; ML trees lacked significant bootstrap support for the divergence between primates, treeshrews, and colugos (fig. S13).

A Bayesian relaxed molecular clock approach with eight fossil constraints estimated the origin of Euarchontoglires at 88.8 million years ago (My), Euarchonta at 87.9 My, and Primatomorpha at 86.2 My (see Fig. 2 and table S4 for 95% credibility intervals). Our divergence dates for Hominoidea/Cercopithecoidea (26.8 My), Anthropoidea (41.7 My), *Lemur/Microcebus* (40.4 My), Strepsirhini (62.1 My), and Primates (79.6 My) were very similar to those estimated from an independent 59.7-kb alignment of the *CFTR* gene region (26) (table S4). The rapid divergence across the basal euarchontan nodes explains why, despite the seven indels and high bootstrap and Bayesian support for Primatomorpha, we were not able to reject the Sundatheria hypothesis on the basis of sequence data alone (Shimodaira-Hasegawa test, $P = 0.065$) (23). We did reject an alliance of treeshrews and primates ($P = 0.047$), despite the single discrepant indel supporting primates + tree shrews. This observation is similar to other findings of incomplete lineage sorting in the common ancestor of rapidly diversifying eutherian clades (27, 28).

The inclusion of nuclear gene sequences from ptilocercid treeshrews allowed us to date the origin of extant treeshrews (Scandentia) to ~63.4 My (Fig. 2 and table S4), near the Cretaceous-Tertiary boundary, concomitant with divergence estimates of many eutherian orders and consistent with the long-fuse model of eutherian diversification (25). This deep divergence between *Ptilocercus* and other scandentians complements profound anatomical and behavioral distinctions that have been documented between these groups (2, 13, 21, 29) and vindicates recent classifications that have separated *Ptilocercus* in a unique family, Ptilocercidae (21, 22). As the sole living representative of a eutherian lineage that diverged in the early Tertiary along with many modern mammalian orders, we suggest that the phylogenetic uniqueness of *Ptilocercus*, combined with its restriction to lowland forest habitats within a relatively limited global range, should render it an important conservation priority in global context.

Because our conclusions imply that colugos, rather than treeshrews, are the most appropriate outgroup for Primates in studying the evolution of adaptive traits, these results may affect the placement of euarchontan fossils and our understanding of primate genomic evolution (3–5). For example, a recent morphological anal-

ysis supporting Sundatheria placed extinct plesiadapiforms in a monophyletic clade with Primates (3), in contrast to Beard (12), who identified plesiadapiforms as members of Dermoptera, within Primatomorpha. Our reanalysis of the data set from (3) that constrains the monophyly of Euprimates and Dermoptera agrees with the placement of plesiadapiforms as the sister group to Euprimates, though this result is only weakly supported (3) (fig. S14). Finally, our results indicate that a draft genome sequence from a colugo is a necessary prerequisite to accurately reconstruct the ancestral primate genome (5).

References and Notes

1. S. Tavaré, C. R. Marshall, O. Will, C. Soligo, R. D. Martin, *Nature* **416**, 726 (2002).
2. E. J. Sargis, *Science* **298**, 1564 (2002).
3. J. I. Bloch, M. T. Silcox, D. M. Boyer, E. J. Sargis, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1159 (2007).
4. M. Goodman, L. I. Grossman, D. E. Wildman, *Trends Genet.* **21**, 511 (2005).
5. E. Pennisi, *Science* **316**, 218 (2007).
6. W. J. Murphy et al., *Science* **294**, 2348 (2001).
7. P. J. Waddell, H. Kishino, R. Ota, *Genome Inform.* **12**, 141 (2001).
8. J. O. Kriegs, G. Churakov, J. Jurka, J. Brosius, J. Schmitz, *Trends Genet.* **23**, 158 (2007).
9. R. D. Martin, *Primate Origins and Evolution* (Princeton Univ. Press, Princeton, NJ, 1990).
10. M. J. Novacek, *Nature* **356**, 121 (1992).
11. J. Shoshani, M. C. McKenna, *Mol. Phylogenet. Evol.* **9**, 572 (1998).
12. K. C. Beard, in *Mammal Phylogeny: Placentals*, F. S. Szalay, M. J. Novacek, M. C. McKenna, Eds. (Springer, New York, 1993), pp. 129–150.
13. E. J. Sargis, *Evol. Anthropol.* **13**, 56 (2004).
14. F.-G. R. Liu et al., *Science* **291**, 1786 (2001).
15. M. S. Springer, M. J. Stanhope, O. Madsen, W. W. de Jong, *Trends Ecol. Evol.* **19**, 430 (2004).
16. R. M. Adkins, R. L. Honeycutt, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10317 (1991).
17. P. J. Waddell, S. Shelley, *Mol. Phylogenet. Evol.* **28**, 197 (2003).
18. J. Schmitz, M. Ohme, B. Suryobroto, H. Zischler, *Mol. Biol. Evol.* **19**, 2308 (2002).
19. O. Madsen et al., *Nature* **409**, 610 (2001).
20. W. J. Murphy et al., *Nature* **409**, 614 (2001).
21. M. C. McKenna, S. K. Bell, *Classification of Mammals Above the Species Level* (Columbia Univ. Press, New York, 1997).
22. K. M. Helgen, in *Mammal Species of the World*, D. E. Wilson, D. M. Reeder, Eds. (John Hopkins Univ. Press, Baltimore, MD, 2005), pp. 104–109.
23. See supporting material on Science Online.
24. D. Huchon et al., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7495 (2007).
25. M. S. Springer, W. J. Murphy, E. Eizirik, S. J. O'Brien, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1056 (2003).
26. M. E. Steiper, N. M. Young, *Mol. Phylogenet. Evol.* **41**, 384 (2006).
27. H. Nishihara, M. Hasegawa, N. Okada, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9929 (2006).
28. W. J. Murphy, T. H. Pringle, T. Crider, M. S. Springer, W. Miller, *Genome Res.* **17**, 413 (2007).
29. L. H. Emmons, *Tupaia: A Field Study of Bornean Tree Shrews* (Univ. of California Press, Berkeley, CA, 2000).
30. This work was supported in part by NSF (grants EF0629849 to W.J.M. and EF0629860 to M.S.S.) and the National Institutes of Health (grant HG02238 to W.M.). We thank A. Jambhekar, T. Crider, A. Wilkerson, V. David, K. Durkin, D. Wilson, L. Grassman Jr., and A. Wilting for technical advice and support and the Broad Institute/Massachusetts Institute of Technology, Baylor College of Medicine–Human Genome Sequencing Center, and Washington University Genome Sequencing Center for access to unpublished sequence data. Sequences from this study have been deposited in GenBank with accession numbers EU142140–EU142251 and EU213052–EU213059.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/792/DC1
Materials and Methods
Figs. S1 to S14
Tables S1 to S5
References

9 July 2007; accepted 1 October 2007
10.1126/science.1147555

A Gene Regulatory Network Subcircuit Drives a Dynamic Pattern of Gene Expression

Joel Smith, Christina Theodoris, Eric H. Davidson*

Early specification of endomesodermal territories in the sea urchin embryo depends on a moving torus of regulatory gene expression. We show how this dynamic patterning function is encoded in a gene regulatory network (GRN) subcircuit that includes the *otx*, *wnt8*, and *blimp1* genes, the cis-regulatory control systems of which have all been experimentally defined. A cis-regulatory reconstruction experiment revealed that *blimp1* autorepression accounts for progressive extinction of expression in the center of the torus, whereas its outward expansion follows reception of the Wnt8 ligand by adjacent cells. GRN circuitry thus controls not only static spatial assignment in development but also dynamic regulatory patterning.

The genomic regulatory code that controls the specification of the future skeletogenic, gut endoderm, and nonskeletogenic mesodermal components of the sea urchin embryo is embodied in a gene regulatory net-

work (GRN). The GRN states the interactions of about 50 genes encoding transcription factors, as determined in an extensive perturbation analysis along with other data (1, 2). The subcircuits of this network control the establishment of tran-

sient regulatory states in the spatial domains of the developing embryo. Here we consider a GRN subcircuit, the function of which is to direct a dynamically expanding ring or torus of regulatory gene transcription early in sea urchin embryogenesis. Transcription of the torus regulatory genes begins at the vegetal pole of the egg in the newly born fourth-cleavage micromeres. These cells give rise to the skeletogenic lineages of the embryo. Transcription of the earliest torus genes starts at about 6 hours after fertilization, then extends to the adjacent ring of mesodermal blastomeres in the early blastula stage (12 hours), and finally encompasses the precursor cells that will generate the gut just before mesenchyme blastula stage (>18 hours) (Fig. 1A). However, within a few hours after these genes are first activated, their expression is extinguished, first in the skeletogenic domain and then in the mesodermal domain.

Determination of the GRN underlying endomesodermal development in the sea urchin embryo (1, 2) has revealed that the key driver of the dynamic torus pattern is the GRN subcircuit shown in Fig. 1B. To understand the operation of this subcircuit, it is important to

note that the cis-regulatory control apparatuses of both *wnt8* and *blimp1* function as AND operators (3); that is, *wnt8* expression requires both β -catenin/TCF and Blimp1 inputs (4) and *blimp1* expression requires both β -catenin/TCF and Otx inputs (table S1). Morpholino-substituted antisense oligonucleotide (MASO) targeting *blimp1* mRNA (5) or *wnt8* mRNA (4) blocks endomesoderm specification.

Expression of the *wnt8* gene illustrates the canonical torus pattern and directly controls its expansion. Several different *wnt* genes are expressed in the sea urchin embryo (6). Although earlier evidence from sea urchin and *Xenopus* indicated that Wnt8 is probably responsible for driving progressive β -catenin nuclearization during cleavage (7–9), we found that Wnt8 is responsible for producing the β -catenin/TCF input, which, according to cis-regulatory analysis, is obligatory for *blimp1* expression (table S1). Thus, MASO repression of Wnt8 expression eliminates 80 to 98% of early *blimp1* expression (fig. S1).

High-resolution measurements of *blimp1* mRNA by quantitative real-time fluorescence polymerase chain reaction (fig. S2, A and B) show that a small amount of *blimp1* mRNA is present maternally; however, there is no maternal *wnt8* mRNA. In the early-cleavage embryo, β -catenin localizes to the nucleus, and by fourth cleavage, β -catenin can be

visualized in the newly born micromere nuclei (7). By fifth cleavage, the *wnt8* gene is activated in the micromeres (Fig. 1C). Because β -catenin/TCF and Blimp1 are the required inputs into the *wnt8* gene, maternal Blimp1 factor must be available, consistent with the evidence that this gene is maternally expressed (fig. S2A). When the *wnt8* gene begins to be transcribed, its response to its own signal transduction system produces a positive feedback circuit between adjacent endomesodermal cells that both produce and receive Wnt8 (1, 4). Otx protein is also nuclearized initially in the micromeres early in cleavage (10), hence it is available ab initio. *blimp1* transcription is activated one cleavage after *wnt8* transcription (Fig. 1C and fig. S2A). Activation must depend on the enhanced level of the β -catenin/TCF input driven by *wnt8* transcription itself. Once both genes are transcribed in the same cells (i.e., from sixth cleavage on), the subcircuit architecture (Fig. 1B) indicates that the patterns of expression of the mutual regulatory partners, *blimp1* and *wnt8*, should be similar. This was confirmed by whole-mount in situ hybridization (WMISH) (Fig. 1C), and their patterns of expression are equally represented in Fig. 1A.

An essential design feature of the relevant *blimp1* cis-regulatory module is that it includes autorepression sites (5). The architecture of the subcircuit in Fig. 1B suggests that autorepres-

Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed. E-mail: davidson@caltech.edu

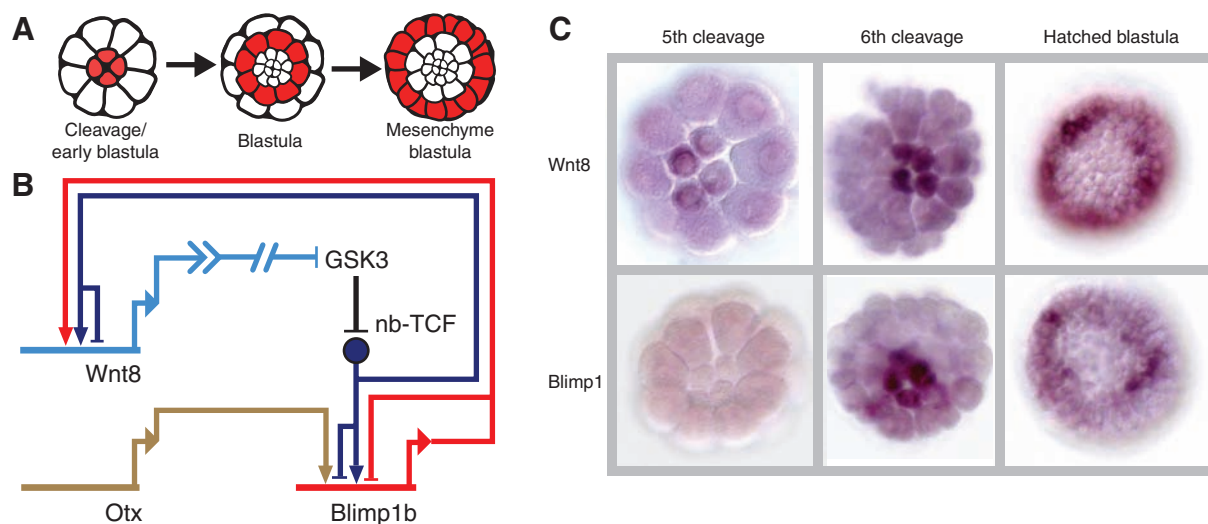


Fig. 1. Moving-torus gene expression pattern. (A) Representation of expression pattern of *blimp1* or *wnt8* genes (red). The innermost cells are skeletogenic micromeres; the red ring in the second drawing shows mesoderm cells (prospective secondary mesenchyme); the outer ring is definitive endoderm. Expression of the *blimp1* gene begins in the micromeres around 6 hours after fertilization and appears in the adjacent tier of mesodermal cells by 12 hours. Soon after, expression disappears from the micromeres. By 18 hours, expression of *blimp1* begins in the adjacent presumptive endoderm lineage and disappears from the mesodermal cells. (B) GRN subcircuit including *otx*, *blimp1*, and *wnt8* genes; *blimp1b* indicates the early isoform of the *blimp1* gene (5); nb-TCF, complex of nuclear β -catenin and TCF transcription factor; GSK3, enzyme normally responsible for

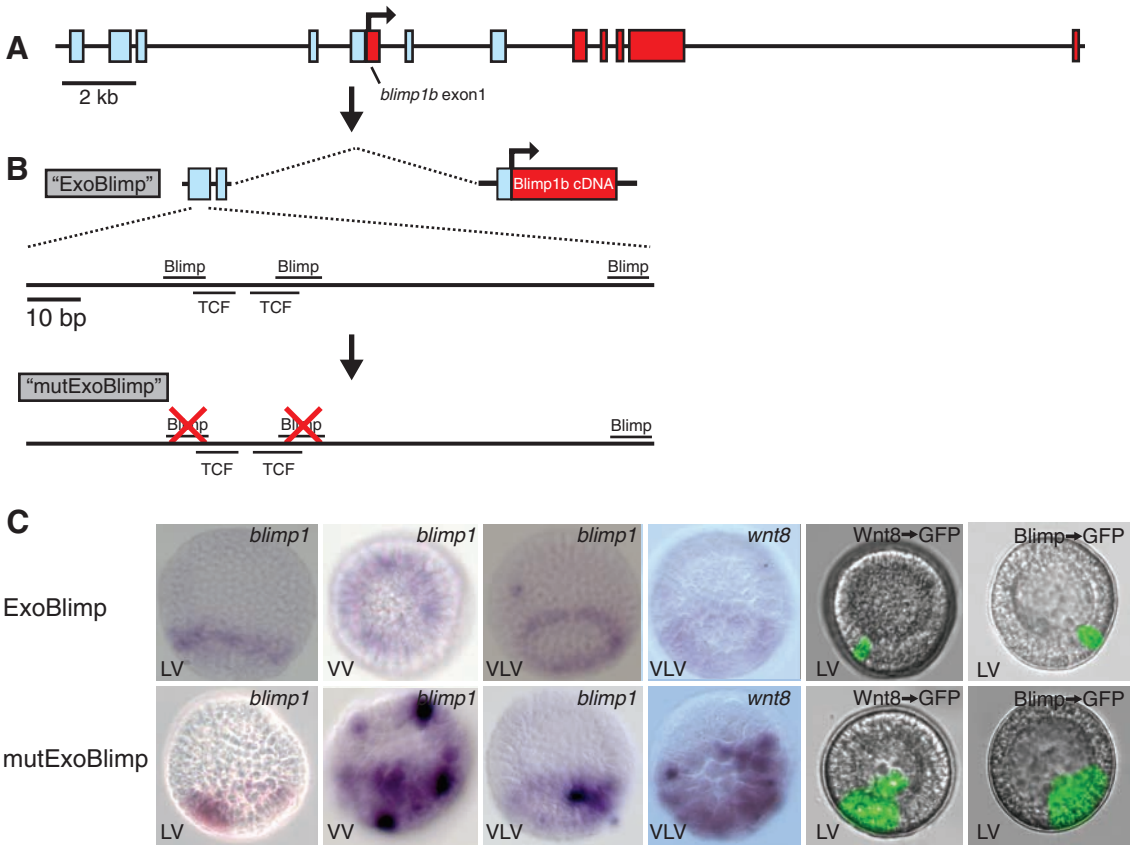
β -catenin clearance, the activity of which is inhibited as a consequence of reception of the Wnt8 signal ligand. In the absence of nuclearized β -catenin, a Groucho/TCF complex forms instead (12) and acts as a dominant repressor at both the *wnt8* and *blimp1* loci (dark blue barred stems). nb-TCF is inhibited from forming by GSK3, the biochemical mechanism of which is symbolized by the solid circle. Positive inputs from Blimp1 and nb-TCF control *wnt8* transcription (4), whereas both nb-TCF and Otx are required for *blimp1* expression; *blimp1* is subject to autorepression via two Blimp1 target sites. (C) Expression of *wnt8* and *blimp1*, visualized by WMISH. By fifth cleavage, *wnt8* transcript is evident in the four micromeres at the vegetal pole. One cleavage later (sixth), *blimp1* transcripts are present in the micromeres. After this, *wnt8* and *blimp1* are expressed in the same territories.

sion of the *blimp1* gene, some hours after its activation, could account for the progressive clearance of both *blimp1* and *wnt8* transcripts from the center of the moving torus of regulatory gene transcription. A series of cis-regulatory reengineering experiments showed that this is indeed the mechanism of clearance. We used a *blimp1* cDNA expression construct that produces normal *blimp1* mRNA under control of the cis-regulatory module responsible for early *blimp1* expression (5) (Fig. 2, A and B). When the cis-regulatory autorepression sites were mutated (Fig. 2B), the construct produced patches of mesodermal *blimp1* transcript lying within the endogenous (mesodermal) *blimp1* clearance zone, whereas the control generated only the normal torus pattern of expression (Fig. 2C, first three columns). We used these constructs to test whether, as predicted, ectopic redeployment of *blimp1* mRNA was sufficient to cause continued expression of *wnt8* in mesodermal territories. Endogenous *wnt8* gene expression monitored by WMISH, as well as a coinjected *wnt8* BAC-GFP (bacterial artificial chromosome–green fluorescent protein) knock-in reporter, produced persistent mesodermal expres-

sion in experimental embryos engineered to express *blimp1* in the mesoderm (Fig. 2C, fourth and fifth columns). [See (11) and fig. S3 for quantitative data from these experiments, including an enhanced GFP mRNA output from the *wnt8* BAC-GFP construct.] According to the architecture of the subcircuit in Fig. 1B, positive spatial input into the *blimp1* gene is provided by β -catenin/TCF (i.e., in response to Wnt8 signaling), because the other positive input, Otx, is continuously available throughout the whole region. To test this, we used a *blimp1* GFP reporter bearing mutated Blimp1 target sites. This construct cannot autorepress, but it displays a normal pattern of expression if the location of *blimp1* transcript is normal (Fig. 2C, sixth column, top). Therefore, the restricted pattern of endogenous β -catenin/TCF, due to the restricted domain of Wnt8 expression, suffices for the restricted spatial expression of the *blimp1* gene. But when the restriction of Wnt8 expression in the mesoderm was relaxed by introduction of ectopic *blimp1* mRNA, the expression of the *blimp1* GFP construct lacking Blimp1 target sites was also relaxed (Fig. 2C, sixth column, bottom).

These experiments show that the cause of progressive vegetal clearance of *wnt8* expression is the restricted localization of the Blimp1 input, which is due entirely to *blimp1* autorepression, as portrayed in the network subcircuit of Fig. 1B. The Fig. 1B subcircuit architecture is directly authenticated at the cis-regulatory level (3, 6, 7) and in this work. Its design ordains its function. It consists of two partially overlapping feedback loops, both of which are subject to an autorepression function, one directly and one indirectly. One loop is signal-mediated: Reception of Wnt8 ligand in recipient cells produces the active β -catenin/TCF transcription factor complex that is required for expression of the *wnt8* gene itself. In the endomesoderm, whether or not there is also an autocrine component, adjacent cells are indeed linked through this signal-driven transcription loop [the “community effect” (1)]. The cis-regulatory system of the *wnt8* gene operates as an AND processor, in the sense that it requires both the Blimp1 and β -catenin/TCF inputs for function. Thus, the requirement for Blimp1 links it obligatorily to the second feedback loop. The

Fig. 2. Experimental demonstration of spatial *blimp1* autorepression. (A) Genomic locus. Red boxes, exons; bent arrow, start of transcription; light blue boxes, non-coding sequence patches displaying high conservation between the sea urchin species *Lytechinus variegatus* and *Strongylocentrotus purpuratus*. Only the exons of the early-expressed (*blimp1b*) isoform are shown (5). (B) Expression constructs. “ExoBlimp” combines three conserved regions that are sufficient to reproduce correct *blimp1* expression and to drive expression of *blimp1* cDNA in the normal pattern of expression. In the “mutExoBlimp” construct, the two Blimp target sites indicated are mutated, destroying the autorepression function; the partially overlapping TCF sites and a third possible Blimp site proximal to the Blimp-TCF cluster remain intact. (C) Expression of genes and constructs indicated in upper right corner of each panel in embryos bearing ExoBlimp (top row) or mutExoBlimp (bottom row). LV, lateral view; VV, vegetal view; VLV, oblique lateral-vegetal view. Stable incorporation of injected constructs is mosaic in sea urchins (18); therefore, not all cells in the center of the ring are uniformly stained in mutExoBlimp embryos. In ExoBlimp embryos, expression of *blimp1*



mRNA from the construct is superimposed on the endogenous torus pattern of *blimp1* expression. Wnt8→GFP denotes a BAC containing the *wnt8* gene that contains a GFP reporter sequence in place of exon 1 of the gene, produced by in vitro recombination (11). Blimp→GFP denotes a construct similar to mutExoBlimp [compare with (B)], except that the *blimp1* cDNA has been replaced by a GFP sequence.

second loop consists of the requirement for Blimp1 as a driver of *wnt8* expression and the reciprocal requirement of β -catenin/TCF for *blimp1* expression. The AND processor of the *blimp1* gene similarly links it into the first loop by its obligatory requirement for the β -catenin/TCF input, but the other partner here is the Otx activator. If both cis-regulatory systems did not include AND gates dependent on the β -catenin/TCF input, the subcircuit would not work. The subcircuit has conditional operating features; that is, its behavior depends on the particular inputs it sees. The TCF input can function either negatively or positively, because (except in cells receiving Wnt8 signal) it binds the transcriptional repressor Groucho (12). This keeps the whole subcircuit quiet in the ectoderm. Its second conditional “off” function is the autorepression of the *blimp1* gene, which from time-course data depends on a certain accumulation of *blimp1* mRNA (and factor) (fig. S2C).

The subcircuit acts to produce the moving torus of gene expression, as summarized in Fig. 3. Measurements of the transcript concentrations indicate that during this phase the *blimp1* gene is producing about 50 transcripts per cell-hour [this is only a few percent of the maximum possible transcription rate (13)]. Blimp1 factor

eventually reaches a level where it acts to repress its own transcription when there could be as many as ~1500 molecules per nucleus, given sea urchin translation rates (14) [more than sufficient for target site occupancy by the typical transcription factor (13), particularly in the small-micromere nuclei]. The Blimp1 factor then disappears from these cells and *wnt8* gene expression turns off as the reinforcing feedback loop is broken. The half-life of *blimp1* transcripts is about 1.5 hours in the micromeres and 2.5 hours in the mesoderm, versus a default average of 3 to 5 hours for all polysomal sea urchin embryo mRNAs (15). Meanwhile, however, the Wnt8 ligand has diffused to the next tier of cells, the future mesoderm (middle tier in Fig. 3). The intercellular diffusion rate of Wnt8 plus the molecules to which it is bound is most unlikely to be rate-limiting, given the very small intercellular distance and the rates that have been assumed for this process by others (16). Upon receipt of the Wnt8 signal, the subcircuit is thereby activated within the mesodermal territory, and the same cycle of events runs its course in this tier of cells, where it operates with very similar kinetics (Fig. 1 and fig. S2C). Subcircuit reactivation in the cells on the inside of the torus of gene expression (by inward signaling) is not observed: Once *blimp1* tran-

scription goes off, it stays off. Blimp1, a SET domain protein, could be silencing its own locus by recruitment of factors such as histone methyltransferases (17). The subsequent disappearance of nuclear β -catenin from the cells within the torus could also result in Groucho repression (7).

The genomic regulatory code is a static linear structure, whereas embryonic development is intrinsically a process driven by dynamically changing regulatory states. Here, we resolved a gene regulatory network subcircuit that combines these aspects in one small apparatus. The subcircuit's kinetics depend on the synthesis and turnover rates of the relevant mRNAs and proteins, as well as on the affinities of the transcription factors for their cis-regulatory target sites (13). What the apparatus does, however, depends on the genomic cis-regulatory sequence of the *blimp1* and *wnt8* genes, where its unique features, its feedback loops, AND gates, and autorepression function are encoded.

References and Notes

1. E. H. Davidson, *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic Press, San Diego, CA, 2006).
2. E. H. Davidson *et al.*, *Science* **295**, 1669 (2002).
3. S. Istrail, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4954 (2005).
4. T. Minokawa, A. H. Wikramanayake, E. H. Davidson, *Dev. Biol.* **288**, 545 (2005).
5. C. B. Livi, E. H. Davidson, *Dev. Biol.* **293**, 513 (2006).
6. J. C. Croce *et al.*, *Dev. Biol.* **300**, 121 (2006).
7. C. Y. Logan, J. R. Miller, M. J. Ferkowicz, D. R. McClay, *Development* **126**, 345 (1999).
8. A. H. Wikramanayake *et al.*, *Genesis* **39**, 194 (2004).
9. S. J. Du, S. M. Purcell, J. L. Christian, L. L. McGrew, R. T. Moon, *Mol. Cell. Biol.* **15**, 2625 (1995).
10. C. K. Chuang, A. H. Wikramanayake, C. A. Mao, X. Li, W. H. Klein, *Dev. Genet.* **19**, 231 (1996).
11. See supporting material on Science Online.
12. R. C. Range, J. M. Venuti, D. R. McClay, *Dev. Biol.* **279**, 252 (2005).
13. H. Bolouri, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9371 (2003).
14. E. H. Davidson, *Gene Activity in Early Development* (Academic Press, Orlando, FL, 1986).
15. C. V. Cabrera, J. J. Lee, J. W. Ellison, R. J. Britten, E. H. Davidson, *J. Mol. Biol.* **174**, 85 (1984).
16. S. Sick, S. Reinker, J. Timmer, T. Schlake, *Science* **314**, 1447 (2006); published online 2 November 2006 (10.1126/science.1130088).
17. I. Gyory, J. Wu, G. Fejer, E. Seto, K. L. Wright, *Nat. Immunol.* **5**, 299 (2004).
18. D. L. Livant, B. R. Hough-Evans, J. G. Moore, R. J. Britten, E. H. Davidson, *Development* **113**, 385 (1991).
19. Supported by a California Institute of Regenerative Medicine fellowship (J.S.) and by NIH grant HD-37105.

Supporting Online Material

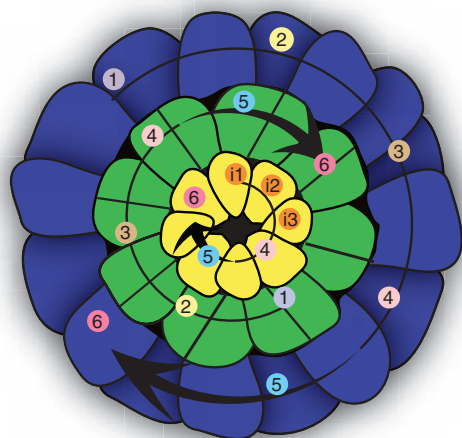
www.sciencemag.org/cgi/content/full/318/5851/794/DC1
Materials and Methods

Figs. S1 to S3

Table S1

References

15 June 2007; accepted 2 October 2007
10.1126/science.1146524



Initial steps (inner tier only)

- 1 Initial State: Groucho-Tcf represses zygotic *blimp1* and *wnt8* expression
- 2 Initial Input: maternal anisotropy stabilizes $n\beta$ -catenin at vegetal pole
- 3 Nuclear β -catenin + maternal Blimp initiate *wnt8* expression; $n\beta$ -catenin + Otx drive zygotic *blimp1* (to step 4)

Initial steps (middle and outer tiers)

- 1 Wnt8 ligand diffuses from inner tier to middle tier or middle tier to outer tier
- 2 $n\beta$ -catenin + Otx drive *blimp1* expression
- 3 $n\beta$ -catenin + Blimp1 drive *wnt8* expression (to step 4)

Common to all tiers

- 4 Wnt8 signaling: among signal-generating cells increases $n\beta$ -catenin; ligand diffuses to neighboring tier (to step 1 in next tier)
- 5 Blimp1 represses self
- 6 In absence of Blimp, *wnt8* turns off

Fig. 3. Summary of mechanism by which the dynamic, concentrically expanding torus of *wnt8* and *blimp1* expression is generated. The drawing shows a seventh-cleavage embryo viewed from the vegetal pole to illustrate the radial concentric organization. However, the events indicated in the numbered key begin at fourth cleavage and extend out to mesenchyme blastula stage (18 to 24 hours).

Telomeric Repeat-Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends

Claus M. Azzalin,^{1,2,3} Patrick Reichenbach,^{1,2,3} Lela Khoraiuli,⁴
Elena Giulotto,⁴ Joachim Lingner^{1,2,3*}

Telomeres, the DNA-protein complexes located at the end of linear eukaryotic chromosomes, are essential for chromosome stability. Until now, telomeres have been considered to be transcriptionally silent. We demonstrate that mammalian telomeres are transcribed into telomeric repeat-containing RNA (TERRA). TERRA molecules are heterogeneous in length, are transcribed from several subtelomeric loci toward chromosome ends, and localize to telomeres. We also show that suppressors with morphogenetic defects in genitalia (SMG) proteins, which are effectors of nonsense-mediated messenger RNA decay, are enriched at telomeres in vivo, negatively regulate TERRA association with chromatin, and protect chromosome ends from telomere loss. Thus, telomeres are actively transcribed into TERRA, and SMG factors represent a molecular link between TERRA regulation and the maintenance of telomere integrity.

Telomeres fulfill essential functions for chromosome stability during mitosis and meiosis, and they equip normal human somatic cells with a cellular clock that determines their replicative lifespan (1–3). Mammalian telomeres comprise tandem arrays of duplex

5'-TTAGGG-3' repeats, with the G-rich strand extending beyond its complement to form an overhang (2, 3). The double-stranded part of telomeres is bound to a multiprotein complex known as shelterin, which prevents telomeres from being recognized and processed as double-stranded breaks (4). The heterochromatic state of telomeres, their gene-less nature, and their ability to silence transcription of experimentally inserted subtelomeric reporter genes, which is known as the telomere position effect (TPE) (5–7), supported the idea that telomeres are transcriptionally silent.

In immortal cells, telomerase promotes telomere length homeostasis by adding telomeric

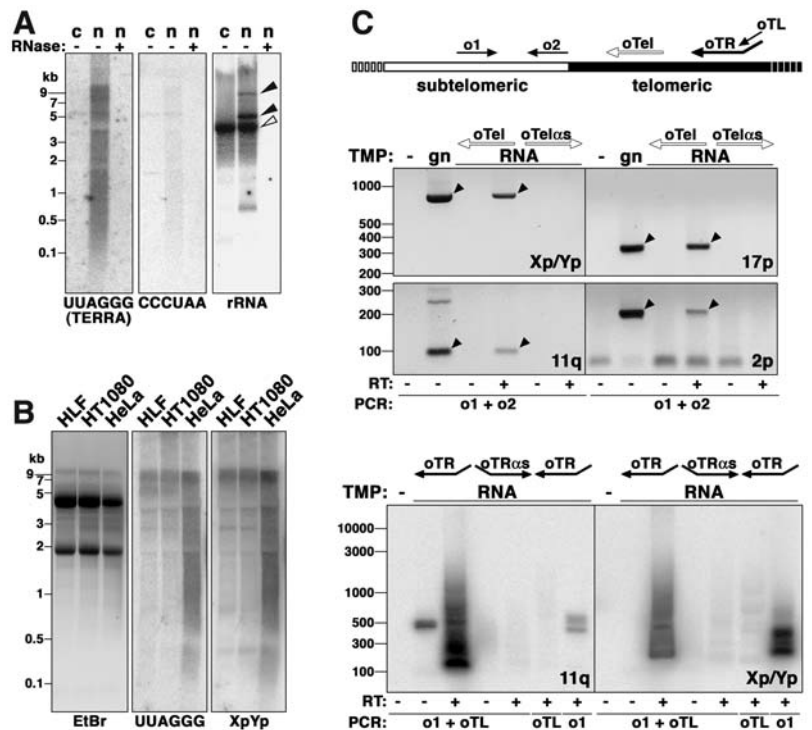
repeats to the 3' ends of chromosomes. The telomerase holoenzyme comprises a reverse transcriptase catalytic subunit, the telomerase RNA moiety, and several accessory factors (2, 3, 8). The *Saccharomyces cerevisiae* Est1 protein binds directly to the telomerase RNA and to the single-stranded telomeric DNA binding protein Cdc13p, which recruits telomerase to chromosome ends in S phase (9–11). Three Est1p-like proteins (EST1A, EST1B, and EST1C) have been identified in humans (12, 13). EST1A and EST1B interact with telomerase, and overexpression of EST1A leads to telomeric fusions and telomere shortening (12, 13). EST1A, EST1B, and EST1C have also been identified as human orthologs of the *Caenorhabditis elegans* SMG proteins SMG6, SMG5, and SMG7, respectively (13–16), which are effectors for nonsense-mediated mRNA decay (NMD) (14). The human NMD core machinery comprises four other SMG polypeptides: UPF1 (SMG2), UPF2 (SMG3), UPF3 (SMG4), and SMG1 (14).

To test whether telomeres are transcriptionally silent, we performed Northern blot analysis of RNA from a human cervical cancer cell line (HeLa) using strand-specific telomeric probes. This approach revealed the existence of TERRA ranging in size from ~100 bases up to at least 9 kilobases (Fig. 1, A and B). The TERRA signal was abolished upon ribonuclease (RNase) treatment, confirming that it originated from RNA (Fig. 1A). TERRA molecules are found exclusively in nuclear fractions and contain UUAGGG repeats. We only detected a faint signal for the complementary CCCUAA repeats (Fig. 1A), suggesting that CCCUAA-containing RNA molecules might exist only at very low levels.

¹Swiss Institute for Experimental Cancer Research (ISREC), CH-1066 Epalinges, Switzerland. ²School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. ³"Frontiers in Genetics" National Centre for Competence in Research (NCCR), CH-1211 Geneva, Switzerland. ⁴University of Pavia, Department of Genetics and Microbiology, 27100 Pavia, Italy.

*To whom correspondence should be addressed. E-mail: joachim.lingner@isrec.ch

Fig. 1. Identification of TERRA. (A) Northern blot analysis of HeLa cytoplasmic (c) and nuclear (n) RNA with strand-specific telomeric and 28S ribosomal RNA (rRNA) probes. Black arrowheads point to nuclear precursors of 28S rRNA (fractionation controls), and the white arrowhead points to the mature 28S rRNA. **(B)** Northern blot analysis of nuclear RNA with a DNA probe from the Xp/Yp subtelomere. Ethidium bromide (EtBr) staining of the gel is shown as loading control. **(C)** HeLa nuclear RNA was reverse transcribed (RT) with oTel or with oTR. Control RT reactions were performed by omitting the RT step or by performing the RT with antisense oligonucleotides (oTelas and oTRas). PCR was performed with the indicated primer pairs and templates (TMP). gn, genomic DNA. PCR products were visualized by EtBr staining (black arrowheads in upper panels) or by hybridization to an oligonucleotide probe (o2) recognizing a specific subtelomeric region distal to o1 (lower panels).



TERRA molecules were detected in different human and rodent cell lines (fig. S1A), indicating conservation of TERRA in mammals.

Hybridization of nuclear RNA blots with a probe derived from the subtelomeric region of human chromosome Xp/Yp generated a TERRA-like hybridization pattern (Fig. 1B). Therefore, human subtelomeres are also transcribed, and TERRA transcription might also start at subtelomeric loci and proceed into the telomeric tract. To test this hypothesis, we reverse transcribed nuclear RNA using a telomeric 5'-(CCCTAA)₅-3' oligonucleotide (oTel) complementary to the putative UUAGGG repeats in TERRA and polymerase chain reaction (PCR)-amplified the derived cDNA using primer pairs located in the subtelomeric regions of chromosomes Xp/Yp, 17p, 11q, and 2p (o1 and o2) (Fig. 1C) (17). The PCR products matched the products stemming from PCR amplifications of genomic DNA (Fig. 1C). Thus, TERRA molecules are transcribed from different telomeres and are composed of subtelomeric-derived RNA and UUAGGG repeats. We also reverse transcribed human nuclear RNA with an oligonucleotide (oTR) containing five 5'-CCCTAA-3' repeats at the 3' end and 20 distinctive nucleotides at the 5' end (oTL) and PCR-amplified the cDNA with the oTL primer in combination with oligonucleotides for subtelomeres of chromosomes 11q and Xp/Yp (o1 in Fig. 1C). The PCR products were heterogeneous in size and positive to Southern hybridization with oligonucleotides (o2 in Fig. 1C) containing sequences that were specific for subtelomeric regions distal to the oligonucleotides used for the PCR. The heterogeneous size distribution of the PCR products is inferred to be due to the presence of many UUAGGG repeats in the 3' region of TERRA and to the annealing of oTR in different registers during reverse transcription.

To analyze the cellular localization of TERRA, we performed RNA fluorescence in situ hybridization (RNA-FISH) experiments. Hybridization

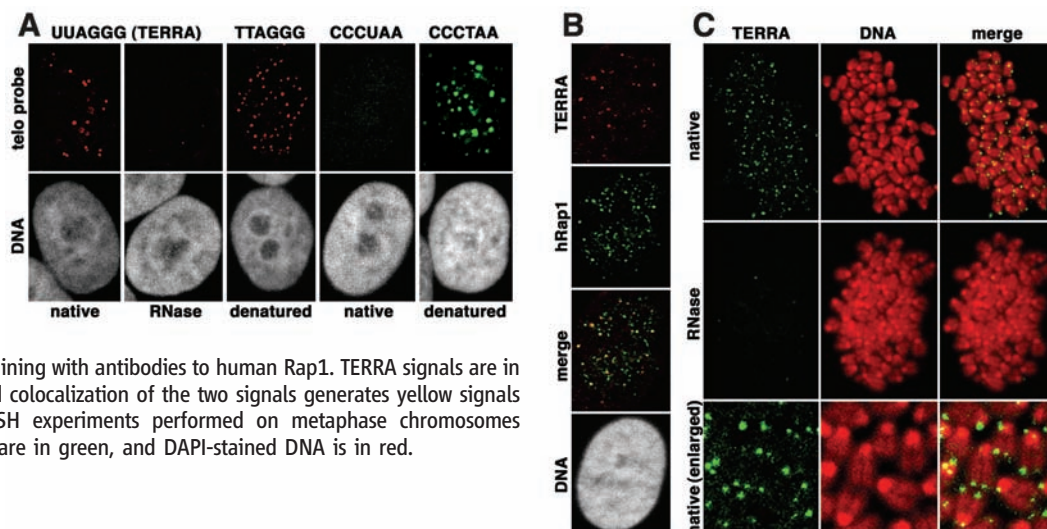
of detergent-extracted nuclei with a fluorescently labeled CCCTAA repeat-containing probe revealed an RNase-sensitive punctate nuclear staining: ~30% of HeLa and primary human lung fibroblasts (HLFs) displayed 3 to 7 TERRA foci, whereas 80 to 100% of human osteosarcoma cells and murine renal cancer cells displayed 20 to 40 foci (Fig. 2A and fig. S1, B and C). Consistent with the Northern blots, our RNA-FISH analysis failed to detect CCCUAA repeat-containing RNA molecules. We then performed RNA-FISH after indirect immunofluorescence, using antibodies against the human shelterin component hRap1 (18). TERRA foci colocalized with Rap1 foci (Fig. 2B and fig. S1D), indicating that TERRA is a component of human telomeric heterochromatin. We also detected TERRA molecules at the ends of metaphase chromosomes from primary mouse embryo fibroblasts (MEFs) (Fig. 2C). The association of TERRA with transcriptionally silent metaphase chromosomes indicates that at least some TERRA remains telomere-associated after its synthesis.

Several SMG proteins are involved in genome stability and telomere maintenance (12, 13, 19, 20). To determine whether SMG proteins regulate TERRA at chromosome ends, we infected HeLa cells with lentiviruses expressing short hairpin RNAs (shRNAs) directed against UPF1, EST1A, SMG1, and UPF2 (fig. S2). Depletion of UPF1, EST1A, or SMG1 induced a substantial increase in the number of cells displaying telomere-associated TERRA foci as well as in the number of foci per positive cell (Fig. 3, A and B, fig. S3, A and B, and fig. S4). A milder increase was observed for UPF2-depleted cells, although the steady-state levels of the NMD substrate BAG1 (21) were similarly increased upon depletion of all tested SMG factors (fig. S3D). The observed accumulation of TERRA in nuclei of SMG-depleted cells was not due to accumulation of cells in different phases of the cell cycle (Fig. 3C and fig. S3C), nor was it the result of substantially increased total TERRA levels (fig. S3D). Further-

more, the approximate half-life for TERRA of 3 hours was not augmented in cells depleted for UPF1 or EST1A (fig. S5). These results suggest that SMG proteins promote the displacement of TERRA from telomeric chromatin and that different SMG proteins might play nonredundant roles in this process. Supporting a direct function for SMG proteins in regulating TERRA at chromosome ends, chromatin fractionation experiments followed by Western blot analysis revealed that all seven SMG proteins are present in chromatin fractions derived from either asynchronous or S-phase HeLa cells (fig. S6A). In addition, chromatin immunoprecipitation experiments with HeLa cell extracts demonstrated that chromatin-bound SMG proteins are several-fold enriched at telomeres in vivo over Alu-repeat sequences (fig. S6, B and C).

We then transfected HeLa cells with plasmids expressing shRNAs against different SMG proteins, which led to more efficient depletion than the lentiviral-mediated depletion used above (fig. S2). Terminal restriction fragment (TRF) analysis in UPF1- and EST1A-depleted cells did not reveal noticeable changes in TRF lengths. Nevertheless, a decrease in the overall intensity of the telomeric smear was seen when the telomeric signal obtained from chromosome ends was normalized to the signal derived from interstitial telomeric repeats (Fig. 4A). Thus, UPF1 and EST1A depletion might lead to the sudden loss of entire telomeric tracts at some chromosome ends. Consistently, DNA-FISH analysis of metaphase chromosomes from UPF1- or EST1A-depleted HeLa cells revealed a five- to sevenfold increase in telomere-free chromosome ends. A twofold increase in associated sister telomeres and a four- to fivefold increase in telomeric fragments were also observed (Fig. 4, B and C, and fig. S7). Similar results were obtained for SMG1-depleted cells, although telomere rearrangements were observed at lower frequencies in cells depleted for UPF2 or EST1C (Fig. 4C and fig. S7). Because plasmid shRNA-mediated depletion of

Fig. 2. Telomeric localization of TERRA. (A) RNA-FISH experiments with strand-specific telomeric DNA probes on HeLa cells. Red and green signals correspond to the probes detecting CCCTAA-CCCUAA repeats or TTAGGG-UUAGGG repeats, respectively. Experiments were performed both in native (to detect RNA) and denaturing (to detect DNA) conditions. 4',6'-diamidino-2-phenylindole (DAPI)-stained nuclei are in gray. (B) RNA-FISH experiments performed on HLF cells after indirect immunostaining with antibodies to human Rap1. TERRA signals are in red, hRap1 signals are in green, and colocalization of the two signals generates yellow signals in the merged panels. (C) RNA-FISH experiments performed on metaphase chromosomes prepared from MEFs. TERRA signals are in green, and DAPI-stained DNA is in red.



UPF1, EST1A, and UPF2 led to similar stabilization of the NMD reporter β -globin NS39 RNA (22) (fig. S2), at least a part of the observed telomeric defects is not an indirect consequence of loss of canonical NMD. Consistently, RNA microarray analysis of cells depleted for UPF1

did not disclose pronounced deregulation of transcripts coding for known telomeric factors, including the six shelterin components (23). Finally, scoring of metaphases depleted for UPF1, EST1A, and SMG1 revealed accumulation of chromosome and chromatid breaks at subtelomeric and

intrachromosomal locations (Fig. 4B), which is consistent with more general roles for UPF1 and SMG1 in genome stability (19, 20) and suggestive of similar roles for EST1A. Mammalian telomeres have genelike properties in that they are actively transcribed into

Fig. 3. SMG proteins promote TERRA dissociation from chromatin. (A) HeLa cells were infected with lentiviruses expressing shRNAs directed against the indicated SMG proteins or empty vector (shEV), and RNA-FISH experiments were performed to detect chromatin-associated TERRA (in red). Numbers indicate the counted TERRA foci for each nucleus. (B) Classification of infected cells according to the number of TERRA foci. 100 to 200 cells were counted for each infection. Means (bars) and SDs (error bars) are from two independent experiments. (C) Fluorescence-activated cell sorting analysis of the same cells as in (A) and (B). PI, propidium iodide.

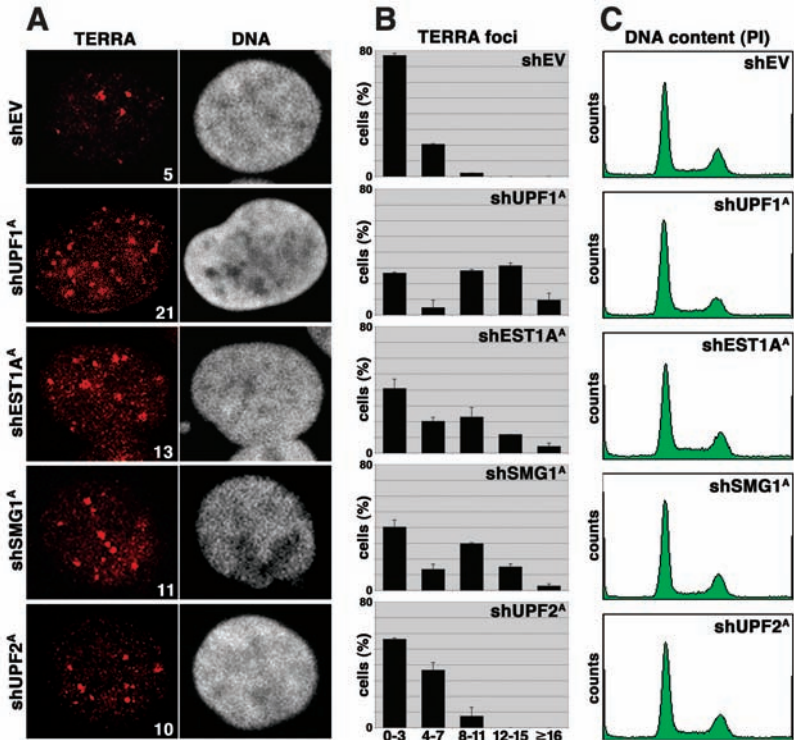
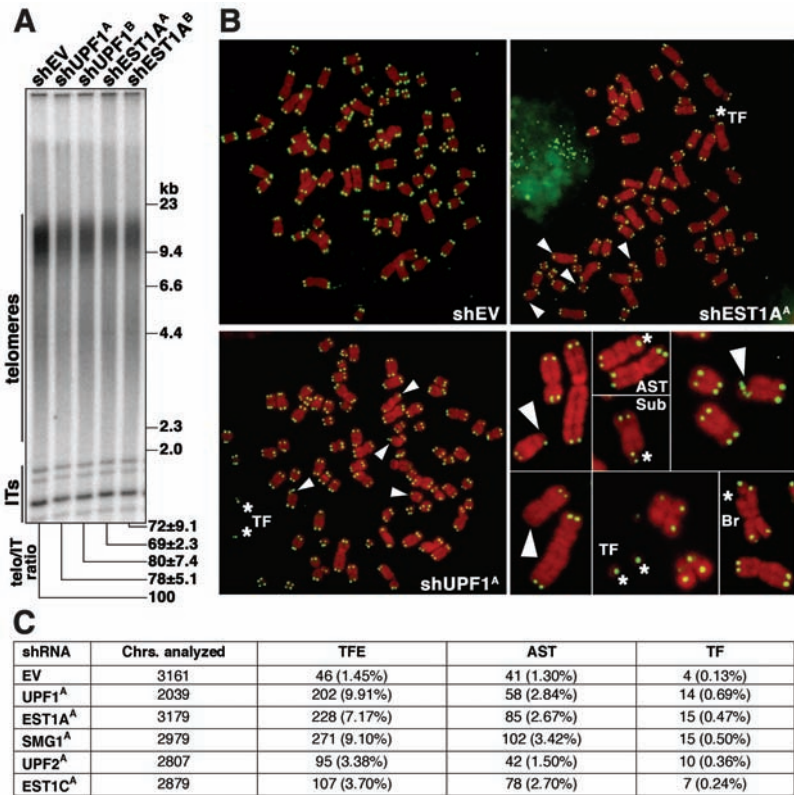


Fig. 4. SMG proteins protect human cells from loss of entire telomere tracts. (A) TRF analysis of genomic DNA prepared from HeLa cells transfected with plasmids expressing shRNAs against UPF1 and EST1A or with empty vector controls. Numbers at the bottom indicate the ratio between the signal associated to the bulk of telomeres and the signal associated to intrachromosomal telomeric repeats (ITs), relative to shEV-transfected cells. Means and SDs are from three independent experiments. (B) Complete metaphases and enlarged examples from telomeric DNA-FISH experiments performed on metaphase chromosomes from HeLa cells treated as in (A). Telomeric sequences are in green, and DAPI-stained chromosomes are in red. Arrowheads point to telomere-free chromosome ends (TFEs). Asterisks indicate telomeric fragments (TFs), associated sister telomeres (ASTs), a subtelomeric chromatid break (Sub), and an intrachromosomal chromatid break (Br). (C) Frequencies of telomeric aberrations observed in SMG-depleted cells.



TERRA molecules. Artificially induced strong transcription of a budding yeast telomere alleviated the TPE exerted by the transcribed telomere and provoked telomere shortening (24), suggesting a role for telomeric RNA in organizing telomere architecture. It is possible that TERRA promotes telomeric heterochromatin assembly by mechanisms similar to the X-chromosome inactivation in females, which is mediated by the long noncoding *Xist* RNA (25), or to the RNA interference-mediated heterochromatinization of the telomeric *dh*-homologous region in fission yeast (26). SMG proteins regulate TERRA at chromosome ends, and UPF1 and EST1A physically interact with DNA polymerase δ and telomerase, respectively (12, 13, 19, 27). We speculate that the telomeric defects induced by SMG-depletion could derive from the loss of coordination between TERRA and key enzymatic activities that assure telomere replication and length homeostasis.

References and Notes

1. T. von Zglinicki, G. Saretzki, J. Ladhoff, F. d'Adda di Fagnaga, S. P. Jackson, *Mech. Ageing Dev.* **126**, 111 (2005).
2. A. Smogorzewska, T. de Lange, *Annu. Rev. Biochem.* **73**, 177 (2004).
3. N. Hug, J. Lingner, *Chromosoma* **115**, 413 (2006).
4. T. de Lange, *Genes Dev.* **19**, 2100 (2005).
5. M. A. Blasco, *Nat. Rev. Genet.* **8**, 299 (2007).
6. J. A. Baur, Y. Zou, J. W. Shay, W. E. Wright, *Science* **292**, 2075 (2001).
7. D. E. Gottschling, O. M. Aparicio, B. L. Billington, V. A. Zakian, *Cell* **63**, 751 (1990).
8. S. B. Cohen *et al.*, *Science* **315**, 1850 (2007).
9. V. Lundblad, J. W. Szostak, *Cell* **57**, 633 (1989).
10. E. Pennock, K. Buckley, V. Lundblad, *Cell* **104**, 387 (2001).
11. T. S. Fisher, A. K. Taggart, V. A. Zakian, *Nat. Struct. Mol. Biol.* **11**, 1198 (2004).
12. B. E. Snow *et al.*, *Curr. Biol.* **13**, 698 (2003).
13. P. Reichenbach *et al.*, *Curr. Biol.* **13**, 568 (2003).
14. I. Behm-Ansmant, E. Izaurralde, *Genes Dev.* **20**, 391 (2006).
15. S. Y. Chiu, G. Serin, O. Ohara, L. E. Maquat, *RNA* **9**, 77 (2003).
16. D. Gatfield, L. Unterholzner, F. D. Ciccarelli, P. Bork, E. Izaurralde, *EMBO J.* **22**, 3960 (2003).
17. D. M. Baird, J. Rowson, D. Wynford-Thomas, D. Kipling, *Nat. Genet.* **33**, 203 (2003).
18. B. Li, S. Oestreich, T. de Lange, *Cell* **101**, 471 (2000).
19. C. M. Azzalin, J. Lingner, *Curr. Biol.* **16**, 433 (2006).
20. K. M. Brumbaugh *et al.*, *Mol. Cell* **14**, 585 (2004).
21. J. Wittmann, E. M. Hol, H. M. Jack, *Mol. Cell. Biol.* **26**, 1272 (2006).
22. R. Thermann *et al.*, *EMBO J.* **17**, 3484 (1998).
23. J. T. Mendell, N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo, H. C. Dietz, *Nat. Genet.* **36**, 1073 (2004).
24. L. L. Sandell, D. E. Gottschling, V. A. Zakian, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12061 (1994).
25. K. Ng, D. Pullirsch, M. Leeb, A. Wutz, *EMBO Rep.* **8**, 34 (2007).
26. J. Kanoh, M. Sadaie, T. Urano, F. Ishikawa, *Curr. Biol.* **15**, 1808 (2005).
27. S. Redon, P. Reichenbach, J. Lingner, *Nucleic Acids Res.*, published online 16 October 2007; 10.1093/nar/gkm724, in press.
28. We thank L. Maquat, M. Blasco, M. Stefanini, J. Lykke-Andersen, O. Mühlemann, H. M. Jack, J. Wittmann, and S. Ohno for reagents and B. Luke for reading the manuscript. J.L.'s laboratory was supported by grants from the Swiss Cancer League, the Swiss National Science Foundation, the Human Frontier Science Program, and the European Union's 6th Framework Programme, and E.G.'s laboratory was supported by grants from the European Commission Euratom, Integrated Project RISC-RAD and the Ministero dell'Università e della Ricerca, Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale (PRIN-2006).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1147182/DC1

Materials and Methods

Figs. S1 to S7

Tables S1 and S2

References

28 June 2007; accepted 17 September 2007

Published online 4 October 2007;

10.1126/science.1147182

Include this information when citing this paper.

A High-Resolution Root Spatiotemporal Map Reveals Dominant Expression Patterns

Siobhan M. Brady,¹ David A. Orlando,^{1,2} Ji-Young Lee,^{1*} Jean Y. Wang,¹ Jeremy Koch,³ José R. Dinneny,¹ Daniel Mace,² Uwe Ohler,^{4,5,6} Philip N. Benfey^{1,6†}

Transcriptional programs that regulate development are exquisitely controlled in space and time. Elucidating these programs that underlie development is essential to understanding the acquisition of cell and tissue identity. We present microarray expression profiles of a high-resolution set of developmental time points within a single *Arabidopsis* root and a comprehensive map of nearly all root cell types. These cell type-specific transcriptional signatures often predict previously unknown cellular functions. A computational pipeline identified dominant expression patterns that demonstrate transcriptional similarity between disparate cell types. Dominant expression patterns along the root's longitudinal axis do not strictly correlate with previously defined developmental zones, and in many cases, we observed expression fluctuation along this axis. Both robust co-regulation of gene expression and potential phasing of gene expression were identified between individual roots. Methods that combine these profiles demonstrate transcriptionally rich and complex programs that define *Arabidopsis* root development in both space and time.

The development of multicellular organisms is regulated by transcriptional networks that act to specify cell types and to provide positional information (1). An understanding of the spatial and temporal control of an organ's transcriptional complexity requires detailed knowledge of its transcriptional states at a resolution specific for cell type and developmental stage. Here, we present, at high resolution, the first microarray-based expression map of a single organ and a profile of nearly all *Arabidopsis* root cell types. With a computational pipeline we developed to analyze this immense data set, we show the output of a complex tran-

scriptional network that underlies root spatiotemporal development, including evidence of fluctuating expression over developmental time and considerable expression variation between individual roots.

The cellular organization of the *Arabidopsis* root reduces the complexity of analyzing its spatiotemporal development. The root consists of 15 cell types that are organized around its radial axis, many of which display rotational symmetry (Fig. 1 and table S1) (2). To a first approximation, the root's rotational symmetry permits the analysis of cell types to be confined to one dimension, the radial axis. The different cell types arise from the

quiescent centre (QC), where initial cells that surround a mitotically less active stem cell niche divide. Cell types are constrained within cell files, so that each new cell division successively displaces an older cell distal to the quiescent centre. Cells undergo division, elongation, and differentiation when they enter the meristematic, elongation, and maturation zones, respectively, along the longitudinal axis (Fig. 1). Because cells are constrained within these files and new cells are born at the root apex, a cell's developmental time line can be tracked along the root's longitudinal axis.

A previous analysis combining fluorescence-activated cell sorting of green fluorescent protein (GFP)-marked cell populations with microarray analysis described expression profiles of five tissue types and three developmental zones in the root (3). These tissue and developmental zone profiles revealed a greater transcriptional complexity than profiles of the organ alone (3). To accurately describe all transcriptional patterns that occur in the root, however, requires a higher-resolution data set that profiles all cell types and developmental stages within an organ. Using the fluorescence cell-sorting expression analysis

¹Department of Biology, Duke University, Durham, NC 27708, USA. ²Program in Computational Biology and Bioinformatics, Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708, USA. ³Providence Glen Drive, Chapel Hill, NC 27514, USA. ⁴Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27708, USA. ⁵Department of Computer Science, Duke University, Durham, NC 27708, USA. ⁶Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708, USA.

*Present address: Boyce Thompson Institute for Plant Research, Tower Road, Ithaca, NY 14853, USA.

†To whom correspondence should be addressed. E-mail: benfey@duke.edu

method (3), we obtained expression profiles of eight new GFP-marked cell populations [S17 (4), S32 (4), COBL9 (5), JO121 (6), S4 (4), SUC2 (7), J2501, and RM1000 (table S1)]. We combined these with 11 previously published experiments [PET111 (8), AGL42 (8), LRC (3), GL2 (3), CORTEX (4), SCR (3), J0571 (3), J2661 (9), APL (4), WOL (3), and S18 (4)] to form a comprehensive data set of 19 experiments (radial data set) that profile expression of 14 nonoverlapping cell types in the *Arabidopsis* root (table S1). To profile developmental stages in the root, we microdissected a single root into 13 sections, with each section encompassing approximately 3 to 5 cells along the longitudinal axis, and we profiled the expression of each by microarray (10). Temporal expression variation in shoot tissue has been demonstrated, and noise in gene expression between genetically identical organisms can determine cell fate (11, 12). Therefore, we also microdissected a second root to assess expression variation between roots in developmental time.

We first asked how expression between individual cell types differed, by using a mixed-model ANOVA (analysis of variance) to determine significant differential expression ($q < 0.001$) (9). We selected 10 marker lines that explore expression variation within the pericycle, and four lines that profile varying developmental stages of xylem and phloem (table S1). Contrary to previous interpretation, the vascular cell types contain the highest number of enriched genes (fig. S1, table S2) (3). Functions of individual cell types have been characterized by genetic or physiological studies, but no comprehensive analysis has described the putative functions of each cell type as defined by its transcriptome. We therefore examined Gene Ontology (GO) term enrichment using the hypergeometric distribution, correcting for multiple hypothesis testing to infer cellular function (13). A number of genes have also been associated with biological processes through microarray expression analysis. We therefore mined the literature for such gene associations and for features that link genes to transcriptional networks through cis-element enrichment (figs. S2 to S4) (10). Two interesting trends were observed for GO term enrichment: First, the majority of enriched terms were associated with individual cell types, and second, those GO terms associated with multiple cell types represent more general biological processes (e.g., peroxisomal activity and membrane localization) (Fig. 2A, and table S3 and fig. S2) (14). As a measure of the reliability of our method in identifying cell type-enriched genes and of our ability to correctly annotate biological processes to a cell type, we compared a list we generated of genes enriched in root hair cells with a previous study that profiled root morphogenesis. In accord with this published report (15), we found root hair cell differentiation genes, kinases, and cell wall-loosening

genes enriched in hair cells ($P = 2.1E^{-4}$, $5.3E^{-6}$, $3.8E^{-4}$) (Fig. 2B). Our analysis further increased the spectrum of biological processes associated with root hair development to include calcium ion transport, vesicle docking during exocytosis, and nicotinamide adenine dinucleotide and/or nicotinamide adenine dinucleotide phosphate ($NAD^+/NADP^+$) activity ($P = 2E^{-4}$, $1.5E^{-4}$, $1.1E^{-4}$, $5E^{-4}$). With this added confidence in our method, we then used enriched gene categories to infer the function of other cell types (table S3 and figs. S2 to S4). Of particular note were xylem cells in the meristematic zone that were enriched for translation initiation and elongation, RNA binding and processing, and mitosis ($P = 7.2E^{-8}$, $5E^{-4}$, $6.1E^{-13}$, $1.8E^{-5}$, $7.3E^{-10}$) (16), (fig. S5). Strong colocalization of the first two enzymes in flavonoid biosynthesis in the cortex has been demonstrated by immunolocalization (17). We provide further evidence for

cortex-enriched flavonoid biosynthesis with enriched expression of four of the five flavonoid biosynthetic enzymes [$P = 6.12E^{-6}$ (table S3)]. Another striking example is the enrichment of genes involved in defense, heat, and oxidative stress responses in the endodermis [$P = 4.3E^{-4}$, $8.97E^{-5}$, $8.23E^{-6}$ (table S3)].

We were also interested in identifying complex spatial expression patterns that are specific to two or more cell types, and therefore, we developed an unsupervised approach to identify dominant, relative expression patterns across cell types and along developmental time (10). If one does not know the a priori expression patterns present in a biological context, there are two methods that can identify them. The first enumerates all possible patterns. However, given the length and complexity of the patterns we are looking for, the number of patterns would be intractably large. Consequently, our pipeline uses

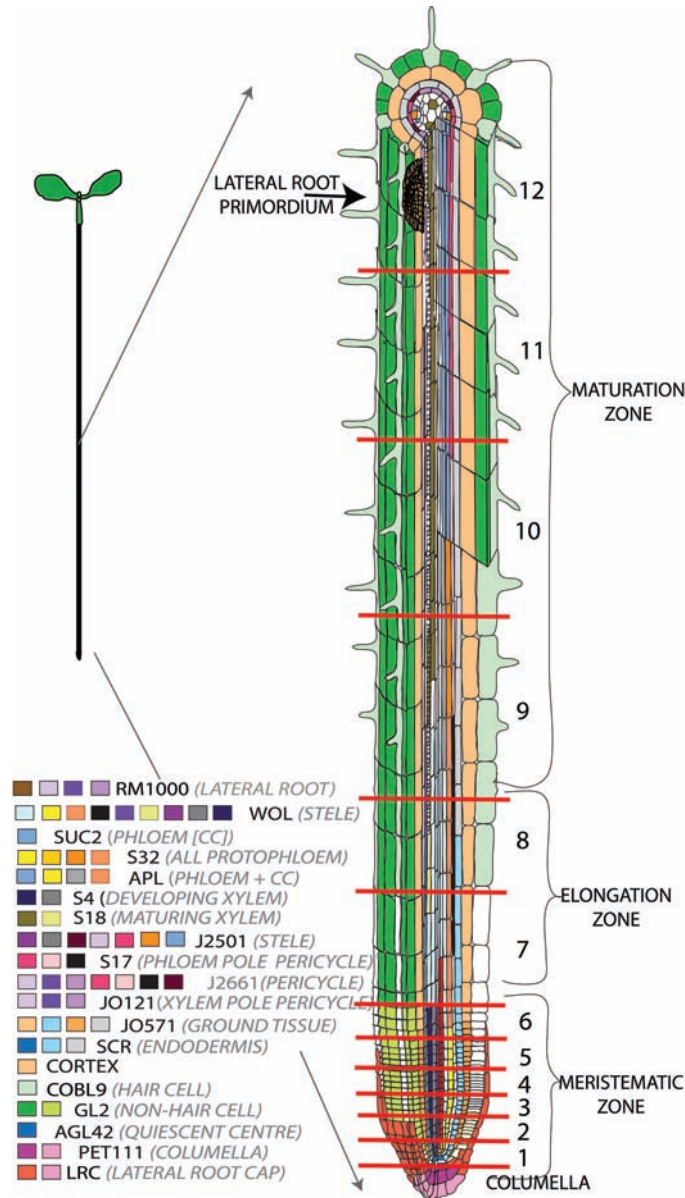


Fig. 1. Microarray expression profiles of 19 fluorescently sorted GFP-marked lines were analyzed (3–9, 23, 24). The colors associated with each marker line reflect the developmental stage and cell types sampled. Thirteen transverse sections were sampled along the root's longitudinal axis (red lines) (10). CC, companion cells.

fuzzy k-means clustering to generate an initial collection of patterns with strong support. Heuristic filters were applied to reduce the collection to a smaller set of distinct expression patterns (10), and 51 distinct, dominant radial and 40 dominant longitudinal patterns resulted (Fig. 2C, Fig. 3A, figs. S17 and S18, and table S4). In comparison, in a previous analysis of five root tissues and three developmental zones, only eight dominant expression patterns were identified using Principle Component Analysis on differentially expressed genes (3). The increased number of identified patterns clearly demonstrates the value of higher-resolution data sets and of the need for more sophisticated computational methods to elucidate underlying transcriptional programs.

We analyzed these patterns to determine the breadth of expression pattern distributions. Seventeen dominant patterns identified from the radial data set show enriched expression in a single cell type (10) (Fig. 2D). In some of the 34 patterns where expression was enriched in multiple cell types, peaks of expression occurred in ontologically or spatially related tissues; in other cases, they were found in cell types that are spatially separated and have no known ontological relationship (Fig. 2, E to G, and fig. S6).

Among the 40 patterns identified from the longitudinal data set (Fig. 3A and fig. S7), the majority showed maximal expression across two or more contiguous root sections (Fig. 3B). This suggests

that, in most cases, the sections we sampled are smaller than the region in which expression is being regulated, which, in turn, indicates that we are at the necessary resolution for detecting differential expression along the longitudinal axis. This also demonstrates a greater complexity of transcriptional programs that underlie previously defined root developmental zones. Development is usually described as a progressive unidirectional process, and it was therefore surprising that 17 of the 40 dominant expression patterns show expression changes that fluctuate over developmental time, presenting multiple peaks of expression (Fig. 3C). Although oscillatory transcriptional mechanisms have been described (18, 19), until recently (20) they had not been described during root development.

To determine the biological relevance of these transcriptional programs, we identified sets of genes that correlated well with the dominant patterns and assessed these lists for biological process enrichment (10) (figs. S8, S19, and S20, tables S5 to S10, and folders S1 and S2). Coexpressed genes assigned to three patterns support transcriptional mechanisms in the regulation of auxin flux across single and multiple cell types. The expression of tryptophan-dependent auxin biosynthetic genes is concordantly enriched in the QC, lateral root primordia, and pericycle [$P = 1.99 \times 10^{-11}$ (Fig. 2E)]; polar auxin transporters and genes that regulate

transporter polarity show enrichment in the columella, root hair cells, and cortex ($P = 1.03 \times 10^{-4}$); and genes that regulate auxin homeostasis are specifically enriched in the columella [$P = 8.82 \times 10^{-4}$ (Fig. 2, F and G)]. In addition to these genes of known function, coexpressed genes of unknown function can be hypothesized to play a role in these processes. In cases where expression is enriched in cell types that have no known ontological or spatial relationship, the biological significance of a shared transcriptional pathway that links these cell types can be inferred. An example is found in hair cells and xylem, which undergo cell wall deposition when they undergo terminal differentiation, and in which a corresponding enrichment of genes involved in cell wall deposition was identified [$P = 1.44 \times 10^{-4}$ (fig. S6)]. Hair cells deposit cell wall materials during hair morphogenesis when the cell rapidly expands, and xylem cells, during the deposition of secondary cell wall. Identifying the functional importance of dominant expression patterns that span multiple cell types should contribute important insights into the network of transcriptional interactions that regulate *Arabidopsis* root development.

Analysis of the genes assigned to longitudinal patterns allowed us to infer and confirm biological processes that occur during root development. Starch-containing amyloplasts in columella cells play a role in gravity sensing and a corresponding

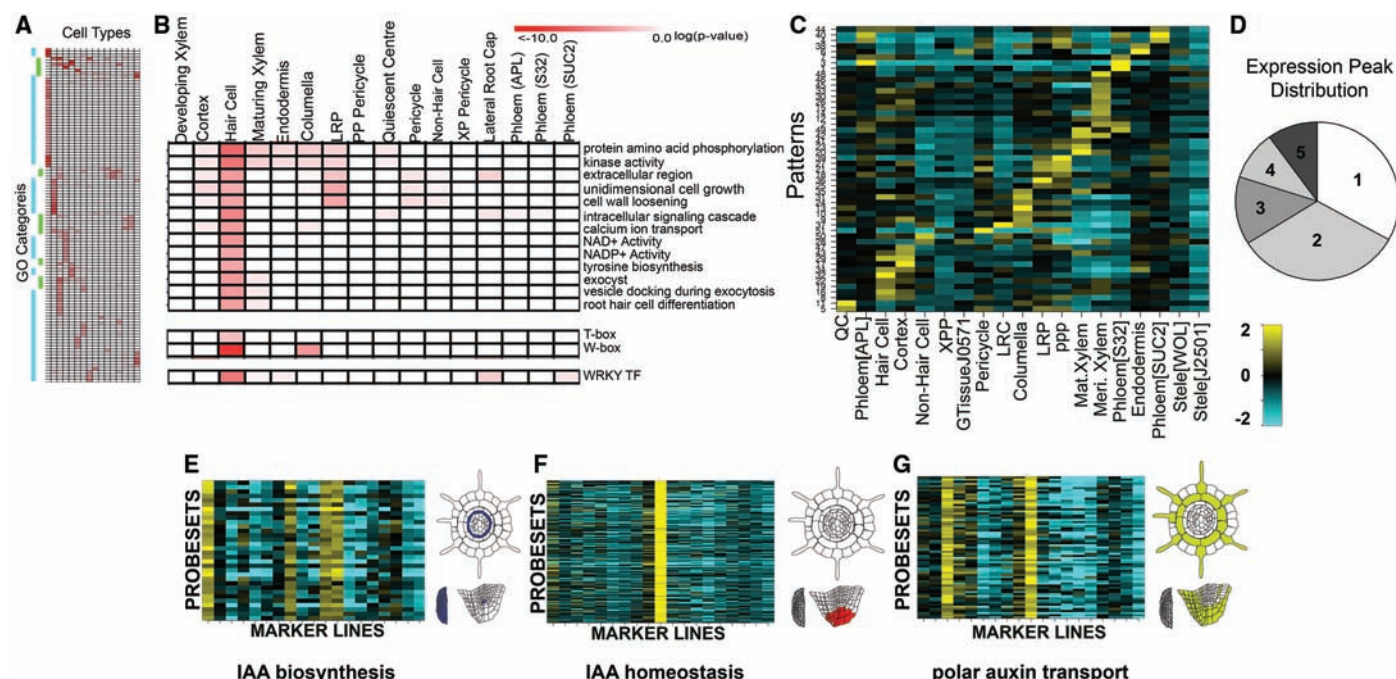


Fig. 2. (A) The majority of enriched GO terms (hierarchically clustered) are associated with individual cell types (blue bar). A smaller number are present across multiple cell types (green bar). (fig. S2) (B) GO category enrichment for hair cells confirms a previous report (15). Enriched cis-elements and an enriched TF family were also identified. (C) From the top 50% of varying probe sets, 51 dominant radial patterns were identified. Pattern expression values were mean-normalized (rows) and \log_2 transformed to yield relative expression indices for each marker line (columns). Marker line order is the same for all figures; see table S1 for marker line abbreviations. (D) Pattern expression peaks were found across one to five cell types. (E to G) Patterns where expression is enriched in single and multiple cell types support transcriptional regulation of auxin flux and synthesis. In all heat maps with

probe sets, expression values were mean-normalized and \log_2 transformed. Expression is false-colored in representations of a root transverse section, a cut-away of a root tip, and in a lateral root primordium. (E) Auxin biosynthetic genes (*CYP79B2*, *CYP79B3*, *SUPERROOT1*, and *SUPERROOT2*) are transcriptionally enriched in the QC, lateral root primordia, pericycle, and phloem-pole pericycle ($P = 1.99 \times 10^{-11}$, pattern 5). All AGI identifiers and TAIR descriptions are found in table S14. (F) Auxin amido-synthases *GH3.6* and *GH3.17* that play a role in auxin homeostasis show enriched expression in the columella, just below the predicted auxin biosynthetic center of the QC ($P = 8.82 \times 10^{-4}$, pattern 13). (G) The expression of the auxin transporter, *PIN-FORMED2*, and auxin transport regulators (*PINOID*, *WAG1*) are enriched in the columella, hair cells, and cortex ($P = 1.03 \times 10^{-4}$, pattern 31).

enrichment of starch catabolism genes was identified here ($P = 2.59E^{-5}$), whereas genes enriched in meristem determinacy and mitosis-specific expression colocalize with the sections containing the QC and just above the dividing initials ($P = 2.22E^{-4}$, $3.79E^{-5}$). By contrast, only one of the patterns that showed fluctuating expression in developmental time correlated with a known biological process—cell cycle activity with expression peaks associated with the root apical meristem and the initiation of presumptive lateral root meristems (Fig. 3E and table S9). For the remaining patterns with fluctuating expression (fig. S9), the majority are enriched in gene functions associated with metabolic or transport activity (table S9).

The pattern-finding described above was performed on expression profiles from a single root because of variability in root growth. To assess the reproducibility of observed expression between replicate roots, we compared the expression of every

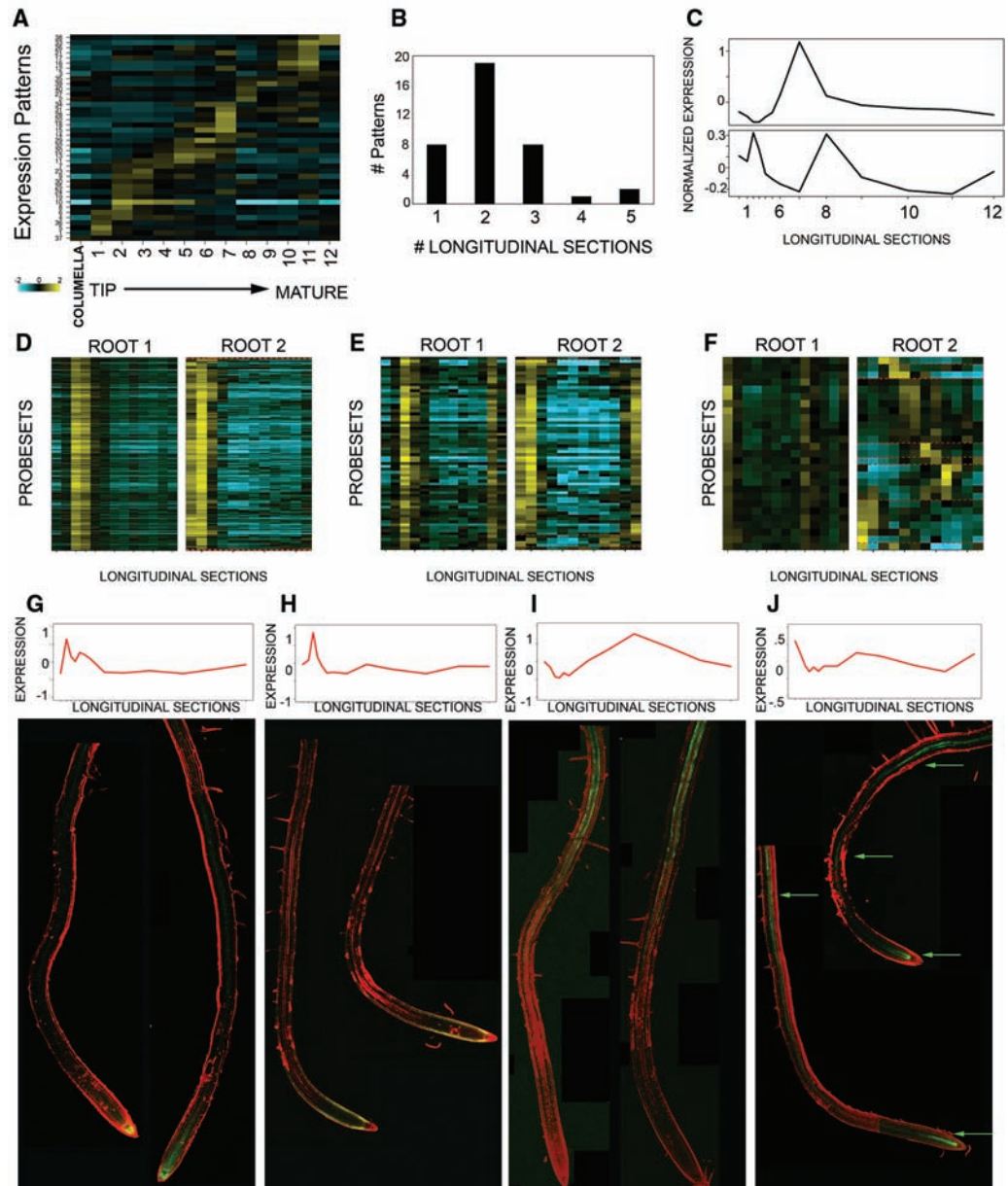
probe set with itself, as well as with randomly selected probe sets. The resulting distributions showed that longitudinal expression is largely reproducible across roots (10) (fig. S10). Next, we asked how robust coexpression is for probe sets assigned to our dominant expression patterns. Methods used to analyze other microarray time series are not suitable because the sections are not identical between roots and the spacing between developmental time points is unknown (21). We therefore determined the extent to which probe sets assigned to dominant patterns in our first root were coexpressed in the second root (10) (Fig. 3, D to F). In the majority of patterns, greater than 90% of the probe sets maintained coexpression (figs. S11 and S12, table S11, and folder S3), which indicated that patterns are quite robust between replicates. In most cases where low coexpression was observed (table S11), subsets of probe sets showed remarkable co-regulation or phasing in several discrete regions along the root. This is

reminiscent of oscillatory expression dependent on the phase of a developmental process (Fig. 3F) (18).

Longitudinal expression profiles were validated by analyzing expression conferred by upstream regulatory regions (4) of representative genes fused to GFP. Expression matched microarray profiles with relative peaks of high expression along the longitudinal axis in all four cases examined as quantified by image analysis (Fig. 3, G to J, and figs. S13 and S14). A gene that displays potentially oscillatory dynamic behavior was also validated (fig. S15), which further demonstrated the reliability of this approach in determining accurate *in vivo* temporal expression programs.

Together, these data sets describe the averaged expression within a cell-type population along portions of the longitudinal axis or among multiple cell types at specific developmental time points. To identify genes with high, relative expression in both space and time, we conditionally added the log₂-

Fig. 3. (A) Forty diverse and distinct patterns were identified from the top 50% of varying probe sets (rows) from sections along the longitudinal axis (columns). (B) Most patterns show an expression increase across two longitudinal sections. (C) Two pattern categories were identified: a single expression peak (top) and fluctuating expression in developmental time (bottom). (D to F) Co-regulation of probe set expression between roots. Patterns were identified from root 1. Root 2 was treated as the replicate. Probe sets are listed in the same order. Red dashed lines indicate co-regulated groups (10). Note that there is no columella section in root 2. (D) Robust co-regulation of expression for a single expression peak (pattern 25). (E) Robust co-regulation of expression for a fluctuating pattern (pattern 19). (F) Extensive between-root variation with multiple co-regulated groups (pattern 37). (G to J) Validation of longitudinal expression patterns with transcriptional GFP fusions. The expression profiles are able to resolve *in vivo* expression at high resolution: *AGAMOUS-LIKE21* (G) expression validates a peak in section 1, whereas *WEREWOLF* (H) shows a peak in section 2; (I) *At4g05170* (pattern 39) confirms the microarray data with a fluorescence peak in the maturation zone. (J) Expression conferred by the *At5g60200* promoter confirms a fluctuating expression profile (pattern 37) with two peaks of expression in the meristematic region and the distal elongation region. This image contains portions distal to the sections dissected for microarray analysis. A third peak of expression is found in this distal maturation zone.



normalized values of the radial and longitudinal dominant expression patterns (Fig. 4, A and B, and folder S4). The resulting high relative peaks at the intersections of expression were visualized in a heat map of the three dimensional (3D) root (Fig. 4B and folder S4). Good correlations were found between expression from the transcriptional fusions and the predicted spatiotemporal expression patterns (Fig. 4B and fig. S16).

Our computational pipeline revealed a number of dominant expression patterns, but the initial data-processing step eliminated genes that were not expressed or that displayed low expression variation. Converting expression to binary (present or absent) values captures different aspects of gene expression discarded when analyzing relative expression. Binary patterns from both data sets were superimposed using the AND logical

operator, so that positive expression in both data sets was required, and the resulting 158 clusters were analyzed (10) (folder S5). By this method, 2018 probe sets were identified as being expressed ubiquitously (Fig. 4C). A number of clusters were found that mark developmental zones (Fig. 4D), cell types at specific developmental stages, and cases where expression was absent in a single cell type throughout developmental time (Fig. 4E).

Expression data associated with our identified patterns has been able to resolve transcriptional programs at high spatial and temporal resolution. We should then be able to use these groups of coexpressed genes to infer transcriptional regulatory modules. Cis-element enrichment within coexpressed gene groups can provide clues as to upstream regulatory transcription factors. We analyzed genes assigned to all patterns for enriched cis-elements

and for the presence of corresponding transcription factors. We were able to infer a MYB-domain transcription factor-regulated auxin biosynthesis module, an auxin response factor (ARF)-regulated module in the columella, and two WRKY-regulated modules in hair cells at different developmental time points (fig. S21). The auxin biosynthesis phenotype conferred by an allele of the *Altered Tryptophan Regulation 1* (*ATR1*) MYB-domain transcription factor supports the validity of our method in inferring these modules (10, 22). More sophisticated algorithms should be able to extend this analysis to identify the complement of transcriptional regulatory modules important for root development.

Together, the data obtained from this high-resolution data set both highlight and revise our view of the rich and complex network of transcriptional programs that underlie cell type and temporal

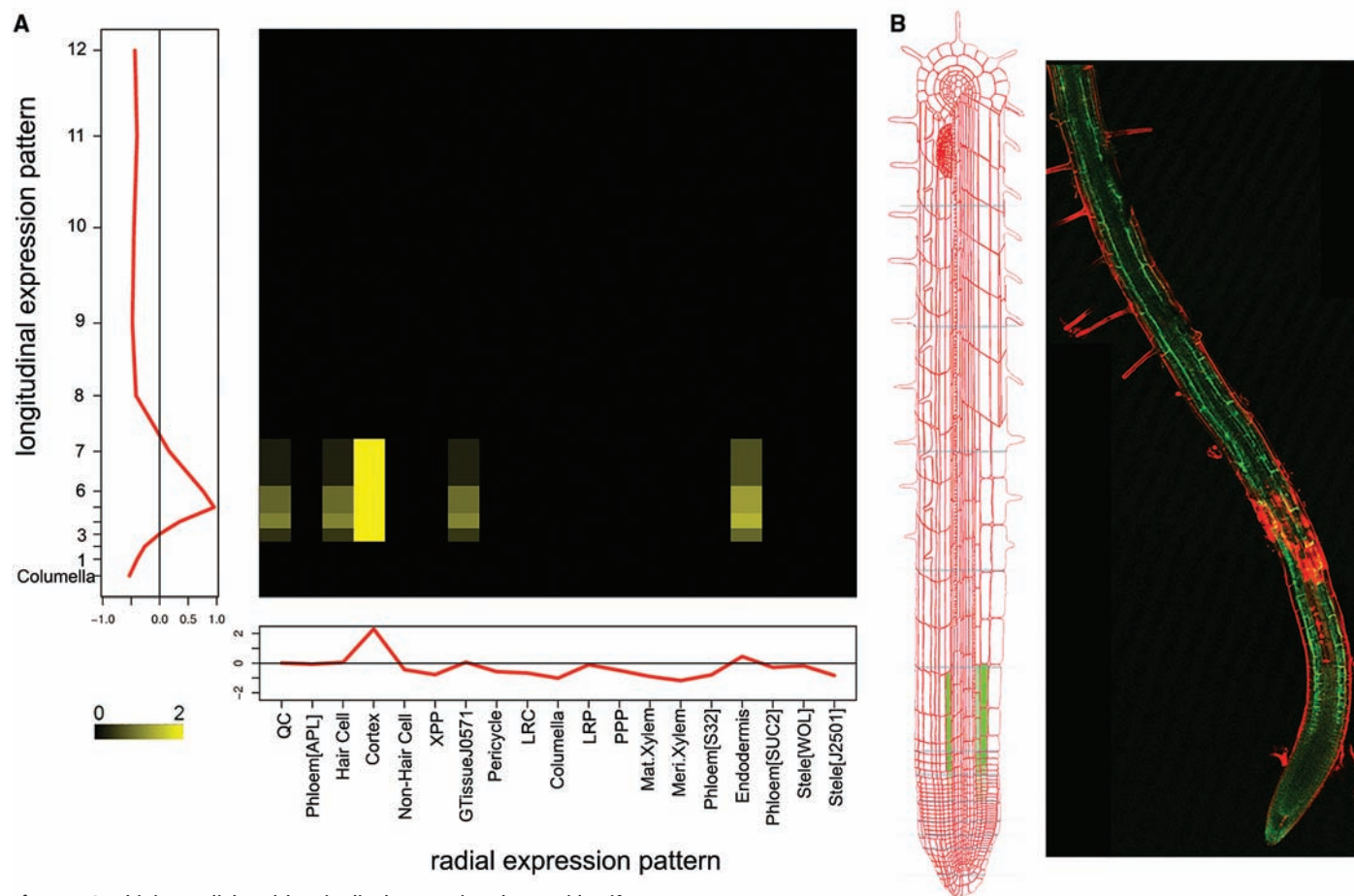


Fig. 4. Combining radial and longitudinal expression data to identify gene expression in space and time using dominant expression patterns (A and B) or binary logic (C to E). (A) Radial (x axis) and longitudinal (y axis) patterns are combined by conditional addition to yield relative enriched peaks of expression by cell type and developmental time. The heat map predicts an enriched relative peak of expression in the cortex and more weakly in the endodermis at the distal region of the meristematic zone and elongation zone. (B) A 3D representation of the conditional addition in (A) (left). The *At3g05150* transcriptional GFP fusion confirms this expression. (C to E) All expression values were converted to binary representation and combined using the logical AND operator. Presence (red) or absence (blue) of expression in the radial data set (x axis) and in the longitudinal data set (y axis) is indicated. Yellow indicates expression found in both data sets. Marker line and longitudinal section order are as in Figs. 2 and 3. (C) Ubiquitous expression in all cell types and developmental stages. (D) Expression only in the maturation zone. (E) Expression in all root zones and cell types except maturing xylem.

aspects of root development. In the future, it will be interesting to compare the spatial and temporal transcriptional complexity that underlies organ development in other multicellular organisms.

References and Notes

1. M. Levine, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4936 (2005).
2. P. N. Benfey, B. Scheres, *Curr. Biol.* **10**, R813 (2000).
3. K. Birnbaum *et al.*, *Science* **302**, 1956 (2003).
4. J.-Y. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6055 (2006).
5. S. M. Brady, S. Song, K. S. Dhugga, J. A. Rafalski, P. N. Benfey, *Plant Physiol.* **143**, 172 (2007).
6. L. Laplace *et al.*, *J. Exp. Bot.* **56**, 2433 (2005).
7. A. Imlau, E. Truernit, N. Sauer, *Plant Cell* **11**, 309 (1999).
8. T. Naway *et al.*, *Plant Cell* **17**, 1908 (2005).
9. M. P. Levesque *et al.*, *PLoS Biol.* **4**, e143 (2006).
10. Materials and methods are available as supporting material on Science Online.
11. E. Yakir, D. Hilman, Y. Harir, R. M. Green, *FEBS J.* **274**, 335 (2007).
12. H. Maamar, A. Raj, D. Dubnau, *Science* **317**, 526 (2007).
13. J. D. Storey, *J. R. Stat. Soc. Ser. B. Method.* **64**, 479 (2002).
14. T. R. O'Connor, C. Dyreson, J. J. Wyrick, *Bioinformatics* **21**, 4411 (2005).
15. M. A. Jones, M. J. Raymond, N. Smirnov, *Plant J.* **45**, 83 (2006).
16. M. Menges, S. M. de Jager, W. Gruissem, J. A. H. Murray, *Plant J.* **41**, 546 (2005).
17. D. Saslow, B. Winkel-Shirley, *Plant J.* **27**, 37 (2001).
18. M.-L. Dequ ant *et al.*, *Science* **314**, 1595 (2006).
19. S. Vanneste *et al.*, *Plant Cell* **17**, 3035 (2005).
20. I. De Smet *et al.*, *Development* **134**, 681 (2007).
21. J. Ernst, G. J. Nau, Z. Bar-Joseph, *Bioinformatics* **21**, i159 (2005).
22. J. L. Celenza *et al.*, *Plant Physiol.* **137**, 253 (2005).
23. K. Birnbaum *et al.*, *Nat. Methods* **2**, 615 (2005).
24. M. Bonke, S. Thitamadee, A. P. Mahonen, M.-T. Hauser, Y. Helariutta, *Nature* **426**, 181 (2003).
25. We thank K. Birnbaum for providing the J2501 data set and B. Parizot and L. Nussbaum for the RM1000 data set, members of the Duke Center for Systems Biology and L. Brady for useful discussions, M. Levesque for assistance with the ANOVA analysis, and T. Vernoux for discussions concerning GO enrichment. We also thank N. Provart, J. Harer, T. Mitchell-Olds, B. Scheres, and members of the Benfey lab for reviewing the manuscript. S.M.B. received partial support from a postdoctoral fellowship from the Natural Science and Engineering Research Council of Canada. J.R.D. is supported by a Ruth Kirschstein National Research Service Award postdoctoral fellowship (NIH). U.O. is an Alfred P. Sloan Fellow in Computational Molecular Biology. This work was funded in large part by an NSF AT2010 grant to P.N.B. (0209754) and by NSF 0618304 to P.N.B. and U.O.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/801/DC1

Materials and Methods

Figs. S1 to S21

Tables S1 to S14

References

Folders S1 to S6

8 June 2007; accepted 2 October 2007

10.1126/science.1146265

Menin Controls Growth of Pancreatic β -Cells in Pregnant Mice and Promotes Gestational Diabetes Mellitus

Satyajit K. Karnik,¹ Hainan Chen,^{1*} Graeme W. McLean,^{1*} Jeremy J. Heit,^{1*} Xueying Gu,¹ Andrew Y. Zhang,¹ Magali Fontaine,² Michael H. Yen,^{1,3} Seung K. Kim^{1,3,†}

During pregnancy, maternal pancreatic islets grow to match dynamic physiological demands, but the mechanisms regulating adaptive islet growth in this setting are poorly understood. Here we show that menin, a protein previously characterized as an endocrine tumor suppressor and transcriptional regulator, controls islet growth in pregnant mice. Pregnancy stimulated proliferation of maternal pancreatic islet β -cells that was accompanied by reduced islet levels of menin and its targets. Transgenic expression of menin in maternal β -cells prevented islet expansion and led to hyperglycemia and impaired glucose tolerance, hallmark features of gestational diabetes. Prolactin, a hormonal regulator of pregnancy, repressed islet menin levels and stimulated β -cell proliferation. These results expand our understanding of mechanisms underlying diabetes pathogenesis and reveal potential targets for therapy in diabetes.

Maternal pancreatic islet expansion in rodents and humans (1–3) suggests that adaptive islet cell growth is a mechanism for ensuring metabolic balance in pregnancy, a physiological state marked by increased insulin demand. Descriptive studies with rats (2, 3) support the hypothesis that proliferation of insulin-secreting islet β -cells is the principal mechanism of β -cell expansion in pregnancy, but the molecular basis of facultative maternal β -cell proliferation is unknown. Moreover, it is unclear if impaired maternal β -cell proliferation leads to reduced insulin levels and gestational diabetes (4).

To investigate the mechanisms controlling maternal islet expansion, we examined β -cell mass in pregnant C57Bl6 mice. We found that maternal β -cell mass increased by twofold (fig. S1A), ac-

commodating increases in maternal body mass (fig. S1B). After parturition, maternal β -cell mass and body mass returned to prepartum levels (fig. S1, A and B). To assess maternal islet cell proliferation, we performed labeling studies with bromodeoxyuridine (BrdU). β -cell proliferation increased in pregnant mice until 15 days postcoitum (dpc) and then declined to prepartum levels (Fig. 1, A to C). Thus, maternal islet β -cell expansion and mass are dynamic in mice.

Hyperplasia of the maternal pituitary (5) and islets in pregnancy is reminiscent of endocrine proliferation in multiple endocrine neoplasia type 1 (MEN1), a human cancer syndrome characterized by synchronous tumors of the pituitary, endocrine pancreas, and parathyroid. Most MEN1 cases result from mutation of *Men1*, whose protein product is menin (6, 7). In mice and humans, mutation and pathological *Men1* loss promote neuroendocrine tumors, including islet β -cell tumors (7, 8). Thus, we postulated that physiological changes in *Men1* expression might regulate facultative maternal β -cell growth in pregnancy. Immunohistology, Western blotting, and real-time reverse transcription polymerase chain reaction (RT-PCR) studies of

maternal islets isolated during gestation revealed that islet levels of *Men1* mRNA and menin decreased in pregnancy (Fig. 1, D to G), then increased to prepartum levels by 1 week after birth of the pups (Fig. 1, F and G). By contrast, we did not detect changes in maternal islet levels of mRNA encoding mixed lineage leukemia-1 (MLL1), a protein that associates with menin (Fig. 1F). Thus, attenuated *Men1* expression corresponded with increased β -cell proliferation in maternal islets.

Menin functions in a histone methyltransferase protein complex containing MLL (9, 10). This complex promotes trimethylation of histone H3 on lysine 4 (H3K4), an epigenetic mark associated with transcriptionally active chromatin. Menin-dependent histone methylation maintains expression of *p27^{Kip1}* and *p18^{INK4C}* (hereafter, *p27* and *p18*), which encode cyclin-dependent kinase (CDK) inhibitors that prevent islet proliferation (11–14). Consistent with our previous findings in islet tumors (11), reduced islet menin levels in pregnancy after 8 dpc were accompanied by reduced *p27* and *p18* mRNA and protein levels (Fig. 1G and fig. S1C) and, as revealed by chromatin immunoprecipitation (ChIP), by decreased levels of menin and trimethyl H3K4 associated with *p27* and *p18* (Fig. 1H). Thus, attenuated menin levels and function reduced pancreatic islet *p27* and *p18* expression in pregnant mice.

To determine whether adaptive maternal β -cell proliferation might require reduced *Men1* expression, we generated mice that permitted conditional *Men1* expression in β -cells. Transgenic mice producing hemagglutinin-tagged menin under control of the tetracycline response element (TRE-*Men1*) (15) were generated and mated with mice expressing the reverse tetracycline trans-activator (rtTA) in β -cells directed by the rat insulin promoter (RIP) (16). In bi-transgenic RIP-rtTA, TRE-*Men1* mice (abbreviated β Men1), administration of doxycycline (Dox) allows rtTA binding to the TRE element and stimulates β -cell expression of *Men1* mRNA and menin protein (fig. S2, A and B). Exposure of RIP-rtTA or TRE-*Men1* single transgenic mice to Dox did not induce changes in menin levels (fig. S2, A

¹Department of Developmental Biology, Stanford University, Stanford, CA 94305, USA. ²Department of Pathology, Stanford University, Stanford, CA 94305, USA. ³Department of Medicine (Oncology Division), Stanford University, Stanford, CA 94305, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: seungkim@stanford.edu

and B). In islets isolated from male or virgin female β Men1 mice continuously exposed to Dox after 10 weeks of age, we detected increased levels of *Men1*, *p27*, and *p18* mRNA. Thus, conditional induction of menin in β -cells stimulated expression of known menin-target genes. However, TRE-*Men1* expression in male or virgin female β Men1 mice did not alter serum insulin or glucose control (fig. S2, C to E), indicating that TRE-*Men1* expression alone did not disrupt β -cell function.

To assess the consequences of menin misexpression during pregnancy, we continuously exposed 10-week-old β Men1 females to Dox for 8

weeks, a period encompassing mating, gestation, and delivery. Western blot and RT-PCR analysis revealed that *Men1* mRNA and menin levels were increased in islets from Dox-exposed pregnant β Men1 females compared to controls, and immunohistology confirmed that menin was increased in islet β -cells from β Men1 mothers (Fig. 2, A to E, and fig. S2F). *p27* and *p18* mRNA and protein in β Men1 islets also increased to levels indistinguishable from those in prepartum islets (Fig. 2A and fig. S2F). By contrast, mRNAs encoding *Insulin1* (*Ins1*), *Insulin2* (*Ins2*), *Glut2*, and *Pdx1*, factors that govern β -cell function, were unchanged in

β Men1 islets (fig. S2G). We measured glucose regulation, islet gene expression, and β -cell growth in female β Men1 mice starting at 10 weeks of age. Unlike age-matched pregnant controls or virgin β Men1 females exposed to Dox, pregnant β Men1 females exposed to Dox and fed ad libitum developed moderate hyperglycemia by 9 dpc, which worsened until delivery (Fig. 2F and fig. S3, A and B). Compared to controls, mean glucose levels in fasted β Men1 mothers on Dox were also significantly higher (Fig. 2G and fig. S3C). Intraperitoneal glucose challenge similarly revealed impaired glucose tolerance in pregnant β Men1 females on Dox

Fig. 1. Dynamic regulation of menin and its target genes during facultative islet growth in pregnant mice. **(A)** Quantification of BrdU incorporation by maternal islet β -cells ($n = 3$ to 6 mice sampled per time point). **(B and C)** Detection of insulin (green) and BrdU (red) in islets from nonpregnant (B) and 17-dpc pregnant mice (C). **(D and E)** Detection of insulin (green) and menin (red) in islets from nonpregnant (D) and 17-dpc pregnant female mice (E). Scale bar, 50 μ m. **(F and G)** Real-time PCR analysis of *Men1* and *Mll* (F) and Western blot analysis of the indicated proteins (G) from isolated maternal islets. **(H)** ChIP studies of menin and trimethylated H3K4 associated with *p18* and *p27* from nonpregnant, pregnant, and postpartum maternal islets ($n = 3$ to 6 mice). Data here and in Figs. 2 to 4 are presented as the means \pm SD. * $P < 0.05$, ** $P < 0.01$. Ab, antibody; IgG, immunoglobulin G; Menin, anti-menin Ab; Methyl, anti-H3K4 Ab; In, input DNA; IP, immunoprecipitate.

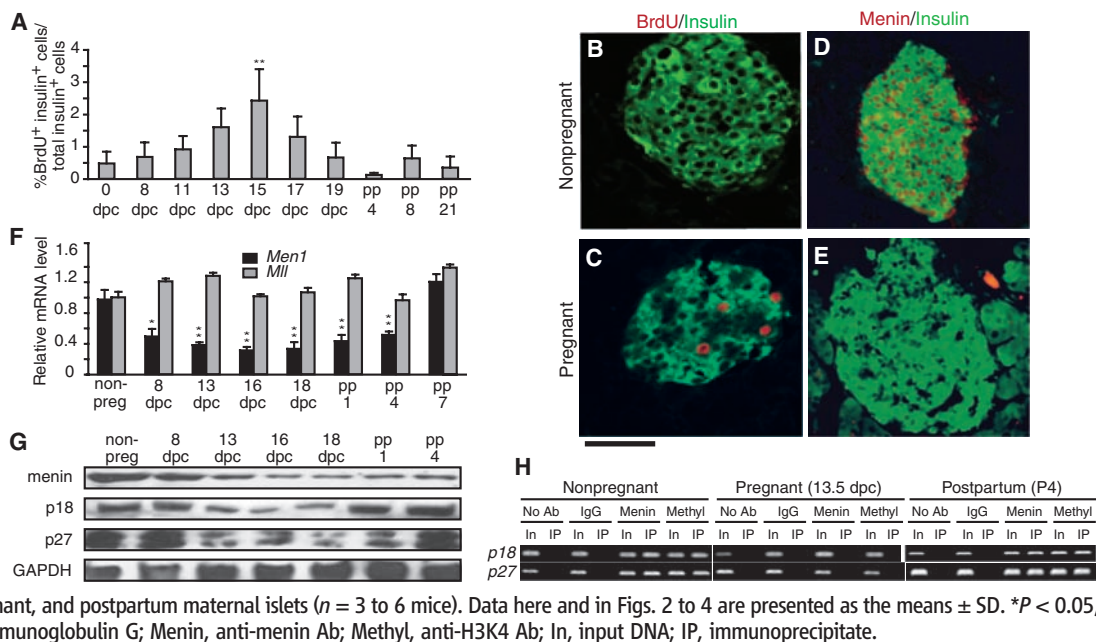
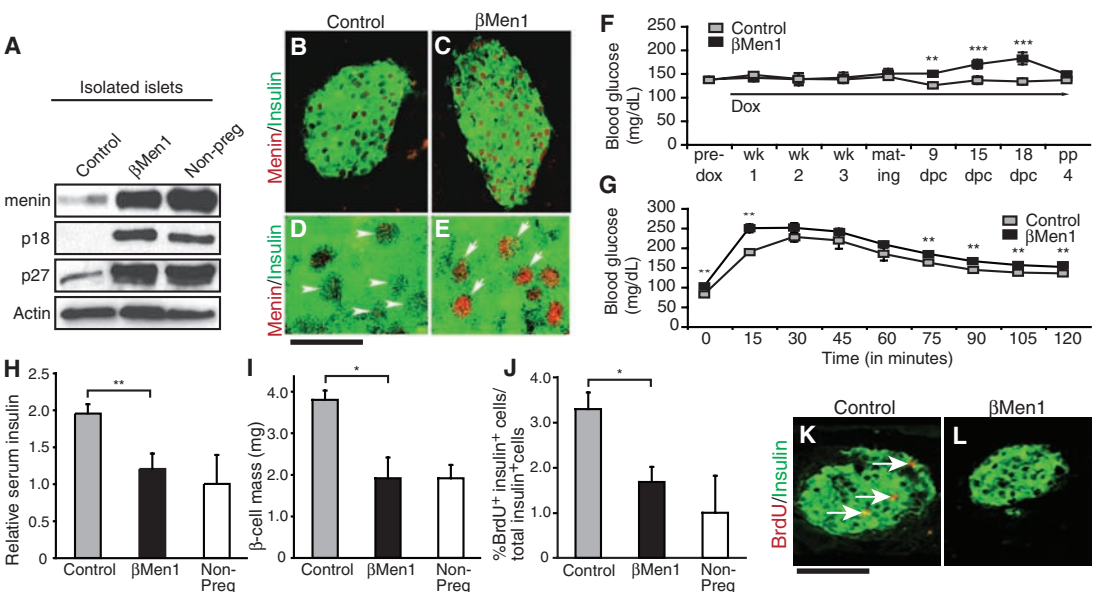
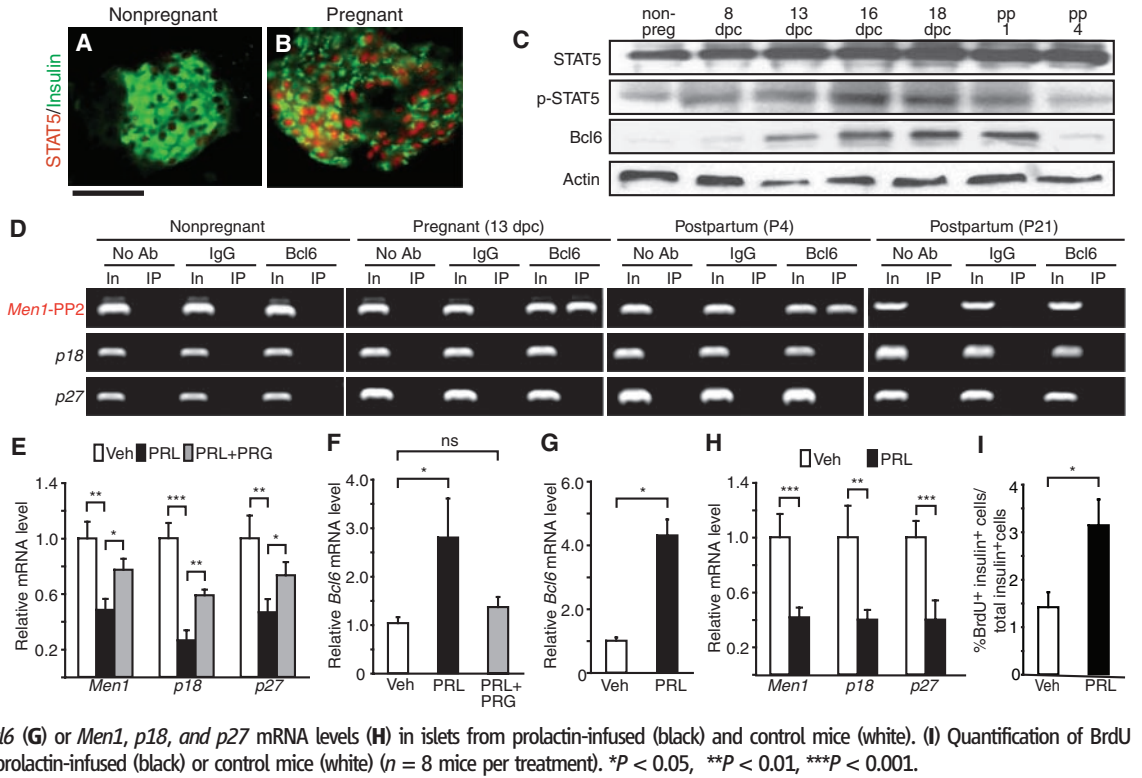


Fig. 2. Phenotypes from misexpression of menin in islets of pregnant mice. **(A)** Western blot analysis of the indicated proteins in islets from pregnant control and pregnant β Men1 mice, both at 17 dpc, and from nonpregnant control mice. **(B to E)** Immunohistologic detection of nuclear menin [red, arrowheads in (D) and arrows in (E)] in islet β -cells (green) from pregnant control (B and D) and pregnant β Men1 (C and E) mice, both at 17 dpc. Scale bar, 50 μ m (B and C) and 20 μ m (D and E). **(F)** Blood glucose levels in β Men1 mice (black) and controls (gray) fed ad libitum and exposed to Dox before, during, and after pregnancy ($n = 10$ to 15 mice per group). **(G)** Intraperitoneal glucose tolerance tests of Dox-exposed pregnant β Men1 (black) and pregnant control mice (gray), both at 16 dpc ($n = 3$ to 6 mice per group). Calculated area under the curve was 25 ± 0.4 area units (β Men1) versus 22 ± 0.7 area units (controls), $P < 0.005$. **(H)** Serum insulin concentrations, **(I)** pancreatic β -cell mass, and **(J)** BrdU incorporation studies in 16- to 17-dpc pregnant β Men1



mice (black), pregnant controls (gray), and nonpregnant controls (white), all administered Dox ($n = 3$ to 6 mice per genotype). **(K and L)** Detection of BrdU (red, arrows) and insulin (green) in pregnant control (K) and pregnant β Men1 mice (L) at 17 dpc. Scale bar, 50 μ m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Fig. 3. Stat5 and Bcl6 regulation of menin in maternal islets. (A and B) Detection of Insulin (green) and Stat5 (red) in islets from nonpregnant (A) and pregnant (B) mice. Scale bar, 50 μ m. (C) Western blot detection of Stat5, phospho-Stat5, and Bcl6 in extracts of maternal islets. (D) ChIP analysis of Bcl6 association with the *Men1*, *p18*, and *p27* loci in islets isolated from nonpregnant, pregnant, postpartum day 4, and postpartum day 21 mice. Primer pair 2 (PP2) for *Men1* has been described previously (7). (E and F) Real-time RT-PCR analysis of *Men1*, *p18*, and *p27* (E) or *Bcl6* mRNA levels (F) in islets exposed to prolactin (black), vehicle (white), or to both prolactin and progesterone (gray) ($n = 3$ independent experiments). Real-time RT-PCR analysis of *Bcl6* (G) or *Men1*, *p18*, and *p27* mRNA levels (H) in islets from prolactin-infused (black) and control mice (white). (I) Quantification of BrdU incorporation by islet β -cells in prolactin-infused (black) or control mice (white) ($n = 8$ mice per treatment). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.



(Fig. 2G). Consistent with these findings, serum insulin levels were reduced in pregnant β Men1 mice (Fig. 2H). Other phenotypes, including body mass, litter size, and average birth weight of pups, were indistinguishable for β Men1 mothers and controls (fig. S4, A to C). In women with gestational diabetes, serum glucose concentration often returns to normal after delivery. Likewise in β Men1 mothers on Dox, blood glucose levels decreased after birth of the pups and were indistinguishable from those of control mothers (Fig. 2F). Thus, β Men1 mice recapitulated features of human gestational diabetes.

To investigate whether impaired insulin production or secretion might underlie hypoinsulinemia in β Men1 mice, we measured insulin content and secretion in isolated islets. Insulin content per islet cell was similar in pregnant β Men1 and control mice (fig. S4D). Likewise, insulin secretion by β Men1 islets after stimulation with glucose or arginine (fig. S4, E and F) was indistinguishable from that of controls. To determine if gestational hypoinsulinemia and hyperglycemia in β Men1 mice reflected impaired β -cell expansion, we assessed the pancreata of pregnant β Men1 and control mice. In Dox-treated pregnant β Men1 mice at the end of gestation, β -cell mass was significantly reduced compared to pregnant controls and appeared indistinguishable from β -cell mass in prepartum β Men1 mice (Fig. 2I). In pregnant β Men1 mice, BrdU studies revealed reduced β -cell proliferation at 17 dpc (Fig. 2, J to L). Thus, gestational reduction of menin, p27, and p18 levels was prevented in β Men1 islets, leading to impaired β -cell expansion. Moreover, we did not detect differences in islet cell apoptosis in β Men1 and control littermate mice (fig. S3G), suggesting that reduced β -cell mass in β Men1 mice did not

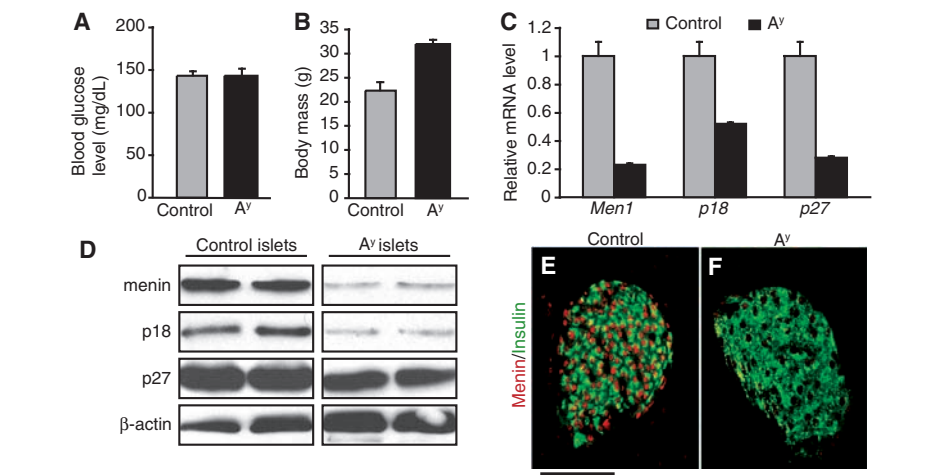


Fig. 4. Dynamic regulation of menin and its target genes during facultative islet growth in obese A^y mice. (A) Blood glucose concentration and (B) body mass in 12-week-old control and A^y littermate mice fed ad libitum. (C) Real-time RT-PCR analysis of *Men1*, *p18*, and *p27* mRNA levels in islets from control or A^y littermate mice. (D) Western blot analysis of the indicated proteins in islets from control and A^y littermate mice. (E and F) Detection of insulin (green) and menin (red) in islet β -cells from control (E) and A^y littermate (F) mice. Scale bar, 50 μ m.

result from cell death. Collectively, these findings suggest that attenuation of maternal islet menin levels permits adaptive β -cell expansion in pregnancy.

Prolactin and placental lactogens are hormonal regulators of pregnancy that stimulate β -cell proliferation in rodent and human islets (17–19), but the molecular basis for their mitogenic effect is unknown. We investigated if prolactin signaling regulates *Men1* in β -cells. Lactogenic hormones stimulate phosphorylation and nuclear accumulation of signal transducer and activator of transcription 5 (STAT5), which induces expression of targets like *Bcl6* (20). In islets from pregnant mice, nuclear

STAT5 accumulated in β -cells (Fig. 3, A and B), and levels of phospho-STAT5 and STAT5 occupancy at consensus STAT5 binding sequences (21) in *Bcl6* increased (Fig. 3C and fig. S5A). *Bcl6* mRNA and protein content increased (Fig. 3C and fig. S5B), and ChIP revealed direct association of Bcl6, a transcriptional repressor (21, 22), with *Men1* during pregnancy (Fig. 3D). Misexpression of Bcl6 in MIN6 cells, a murine insulinoma-derived β -cell line responsive to normal growth cues (11, 23), was sufficient to reduce expression of endogenous *Men1*, *p18*, and *p27* mRNA (fig. S5C). Moreover, *Bcl6* expression reduced transcription of *Men1*-

luciferase reporters harboring consensus Bcl6 binding sequences (fig. S5, D and E). Thus, Bcl6 directly associated with and repressed *Men1* transcription in β -cells. Previous studies have shown that steroids like progesterone and dexamethasone can inhibit the mitogenic effects of prolactin on β -cells (24), but the underlying mechanism is unknown. Simultaneous exposure of isolated mouse islets to prolactin and progesterone attenuated changes in *Men1*, *p18*, *p27*, and *Bcl6* expression provoked by prolactin alone (Fig. 3, E and F). Thus, multiple hormonal inputs likely regulate β -cell *Men1* expression. Bcl6-dependent changes in *Men1*, *p18*, and *p27* expression provoked by prolactin in MIN6 cells or in cultured human islets (figs. S6 to S8) corroborated these findings and showed that *Men1* regulation by lactogen signaling is evolutionarily conserved.

To test if lactogen signaling was sufficient to reduce *Men1* expression and increase β -cell proliferation in vivo, we transplanted mice with osmotic micropumps to deliver prolactin for 6 days (25). Compared with islets from vehicle-infused controls, islets from prolactin-infused mice had a fourfold increase in *Bcl6* expression, a 50% reduction of *Men1*, *p18*, and *p27* mRNA, and a 2.5-fold increase of BrdU incorporation by β -cells (Fig. 3, G to I). Thus, short-term prolactin infusion was sufficient to reduce *Men1* expression in vivo and to stimulate proliferation of adult islet β -cells. Additional studies are needed to determine if lactogenic hormone regulation of *Men1* governs other features that affect β -cell expansion, such as β -cell size and survival (26).

To determine if menin might regulate adaptive β -cell expansion in obesity, another common physiological state that stimulates adaptive islet expansion, we measured islet menin levels in A^y mice, a well-characterized model of hyperphagic obesity [reviewed in (27)]. At 3 months, when A^y mice are obese but normoglycemic (Fig. 4, A and B), A^y islet levels of *Men1* mRNA, menin, and *p27* and *p18* mRNA and protein were reduced compared to islets from wild-type controls (Fig. 4, C to E). These results suggest that in obesity, menin attenuation regulates adaptive β -cell proliferation.

Studies of endocrine neoplasias in MEN1 syndrome and other cancers have defined menin roles solely in the context of tumor pathogenesis. Our work expands this view, showing that menin functions as a physiological regulator of adaptive β -cell expansion in pregnancy and possibly other common states linked to type 2 diabetes, such as obesity. We speculate that menin may integrate β -cell growth signals in physiological islet expansion, controlling dynamic histone modifications that govern β -cell fate and proliferation. Menin-independent control of maternal β -cell expansion is not excluded by our study, and investigating the role of other islet tumor suppressors, like von Hippel–Lindau protein (28), in physiological β -cell expansion might be fruitful. Our finding that *Men1* expression is regulated by prolactin and progesterone raises the possibility that defects in signaling pathways regulated by lactogenic or steroid hormones might underlie specific forms of type 2 diabetes, including gestational diabetes, and endocrine neoplasias linked to

Men1 inactivation, including carcinoid and insulinoma (7). Our work also suggests that manipulation of regulators, cofactors, and targets of menin might be a therapeutic strategy for expanding functional pancreatic islets in diabetes.

References and Notes

1. F. A. Van Assche, L. Aerts, F. De Prins, *Br. J. Obstet. Gynaecol.* **85**, 818 (1978).
2. J. A. Parsons, T. C. Brelje, R. L. Sorenson, *Endocrinology* **130**, 1459 (1992).
3. R. L. Sorenson, T. C. Brelje, *Horm. Metab. Res.* **29**, 301 (1997).
4. J. Boloker, S. J. Gertz, R. A. Simmons, *Diabetes* **51**, 1499 (2002).
5. E. Horvath, K. Kovacs, B. W. Scheithauer, *Pituitary* **1**, 169 (1999).
6. C. Larsson, B. Skogseid, K. Oberg, Y. Nakamura, M. Nordenskjold, *Nature* **332**, 85 (1988).
7. S. K. Agarwal *et al.*, *Ann. N. Y. Acad. Sci.* **1014**, 189 (2004).
8. J. J. Heit, S. K. Karnik, S. K. Kim, *Annu. Rev. Cell Dev. Biol.* **22**, 311 (2006).
9. C. M. Hughes *et al.*, *Mol. Cell* **13**, 587 (2004).
10. A. Yokoyama *et al.*, *Mol. Cell. Biol.* **24**, 5639 (2004).
11. S. K. Karnik *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14659 (2005).
12. T. A. Milne *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 749 (2005).
13. R. W. Schnepf *et al.*, *Cancer Res.* **66**, 5707 (2006).
14. D. S. Franklin, V. L. Godfrey, D. A. O'Brien, C. Deng, Y. Xiong, *Mol. Cell. Biol.* **20**, 6147 (2000).
15. Materials and methods are available as supporting material on Science Online.
16. J. J. Heit *et al.*, *Nature* **443**, 345 (2006).
17. T. C. Brelje *et al.*, *Endocrinology* **132**, 879 (1993).
18. J. H. Nielsen, *Endocrinology* **110**, 600 (1982).
19. L. Labriola *et al.*, *Mol. Cell. Endocrinol.* **264**, 16 (2007).
20. F. A. Scheeren *et al.*, *Nat. Immunol.* **6**, 303 (2005).
21. C. C. Chang, B. H. Ye, R. S. Chaganti, R. Dalla-Favera, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6947 (1996).
22. V. L. Seyfert, D. Allman, Y. He, L. M. Staudt, *Oncogene* **12**, 2331 (1996).
23. J. Miyazaki *et al.*, *Endocrinology* **127**, 126 (1990).
24. A. J. Weinhaus, N. V. Bhagoo, T. C. Brelje, R. L. Sorenson, *Endocrinology* **141**, 1384 (2000).
25. T. Shingo *et al.*, *Science* **299**, 117 (2003).
26. L. Scaglia, F. E. Smith, S. Bonner-Weir, *Endocrinology* **136**, 5461 (1995).
27. T. T. Yen *et al.*, *FASEB J.* **8**, 479 (1994).
28. J. C. Yao, *Best Pract. Res. Clin. Endocrinol. Metab.* **21**, 163 (2007).
29. We thank A. Yokoyama, G. Barsh, X. Hua, J. Hess, B. Cairns, L. Raetzman, A. Levine, and D. Bauer for helpful discussions and advice and members of the Kim laboratory for comments on the manuscript. Supported by a Kirschstein Postdoctoral Fellowship (S.K.K.); the Stanford Regenerative Medicine Training Program (H.C.); the Stanford Medical Scientist Training Program (J.J.H.); the American Diabetes Association (A.Y.Z. and J.J.H.); and NIH grant T32DK007217-32 (M.H.Y.). Work in the Kim laboratory was supported by a Stanford Cancer Council award, a gift from Raymond and Beverly Sackler, the Snyder Foundation, the NIH, the Juvenile Diabetes Research Foundation, and the Larry L. Hillblom Foundation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/806/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S3
References

21 June 2007; accepted 5 October 2007
10.1126/science.1146812

Ordered Phosphorylation Governs Oscillation of a Three-Protein Circadian Clock

Michael J. Rust,^{1*} Joseph S. Markson,^{1,2*} William S. Lane,³ Daniel S. Fisher,⁴ Erin K. O'Shea^{1†}

The simple circadian oscillator found in cyanobacteria can be reconstituted in vitro using three proteins—KaiA, KaiB, and KaiC. The total phosphorylation level of KaiC oscillates with a circadian period, but the mechanism underlying its sustained oscillation remains unclear. We have shown that four forms of KaiC differing in their phosphorylation state appear in an ordered pattern arising from the intrinsic autokinase and autophosphatase rates of KaiC and their modulation by KaiA. Kinetic and biochemical data indicate that one of these phosphoforms inhibits the activity of KaiA through interaction with KaiB, providing the crucial feedback that sustains oscillation. A mathematical model constrained by experimental data quantitatively reproduces the circadian period and the distinctive dynamics of the four phosphoforms.

Circadian clocks coordinate metabolism and behavior with diurnal cycles in the environment (1). These clocks traditionally have been understood as transcriptional feedback oscilla-

tors in which clock genes repress their own synthesis, creating negative feedback that drives oscillation (1). However, pioneering work by Kondo and colleagues has shown that the circadian clock of the cyanobacterium *Synechococcus elongatus* requires neither transcription nor translation (2), and circadian oscillations can be reconstituted in vitro using only three proteins: KaiA, KaiB, and KaiC (3).

KaiC is a hexameric enzyme (4) that can autophosphorylate (5) and autodephosphorylate (6) at both serine 431 (S431) and threonine 432 (T432) (7, 8). The dimeric KaiA (9, 10) enhances the autophosphorylation of KaiC (11), whereas KaiB antagonizes the activity of KaiA (11–13). In the absence of KaiA, KaiC fully dephospho-

¹Howard Hughes Medical Institute, Faculty of Arts and Sciences Center for Systems Biology, Departments of Molecular and Cellular Biology and of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. ²Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA. ³Microchemistry and Proteomics Analysis Facility, Faculty of Arts and Sciences Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA. ⁴Department of Applied Physics, Stanford University, Stanford, CA 94305, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: erin_oshea@harvard.edu

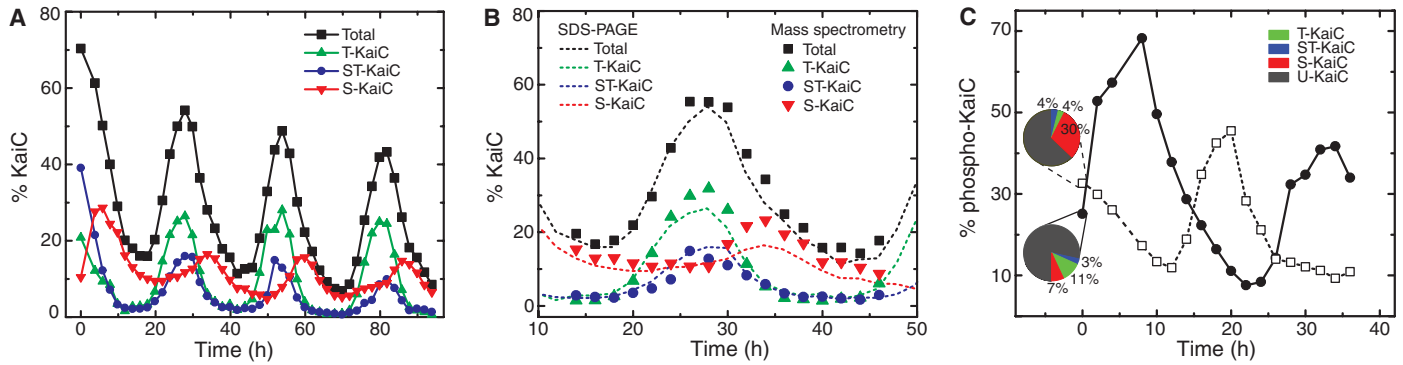


Fig. 1. Phosphorylation of KaiC is cyclically ordered. **(A)** Decomposition of total KaiC phosphorylation ("Total") into its constituent phosphoforms, measured by SDS-PAGE (used throughout this study unless noted otherwise). The percentage of U-KaiC (not shown) is equal to 100% – Total. See also fig. S10. **(B)** Comparison of phosphoform distributions measured by SDS-PAGE [dotted lines, from (A)] and by mass spectrometry (solid symbols). **(C)** The initial phosphoform distribution of KaiC determines the subsequent dynamics. We prepared KaiC enriched in T-KaiC (solid circles) by incubating

unphosphorylated KaiC with epitope-tagged KaiA for 2.25 hours and then removing KaiA by immunoprecipitation. We prepared KaiC enriched in S-KaiC (open squares) by incubating unphosphorylated KaiC with epitope-tagged KaiA for 18 hours, removing KaiA, and then allowing KaiC to autodephosphorylate for 5.5 hours. In both cases, circadian oscillations were then initiated by adding KaiB, incubating for 1.5 hours, then reintroducing KaiA (28). Pie charts show the KaiC phosphoform distribution at the time of KaiA re-addition.

rylates (9). Complexes form between the Kai proteins (14), and the relative proportions of the KaiC-containing complexes oscillate (9, 15).

The amount of phosphorylated KaiC oscillates with a circadian period (11). However, the total level of phosphorylation cannot be the only dynamical variable controlling the oscillator because it traverses the same value twice each day, but each time in a different direction (increasing during the subjective day and decreasing during the subjective night). Previous mathematical models have treated both phosphorylation sites as functionally equivalent (16–24) and have proposed additional dynamical variables arising from persistent conformational changes (18, 20–24) or long-lived heterocomplexes (16, 17); we hypothesized that additional variables could be found by examining the pattern of multisite phosphorylation of KaiC during the circadian cycle.

We measured the time dependence of phosphorylation at S431 and T432 by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 1, A and B) and mass spectrometry (Fig. 1B), quantifying the four possible phosphorylation states: unphosphorylated (U-KaiC), phosphorylated only on S431 (S-KaiC), phosphorylated only on T432 (T-KaiC), and phosphorylated on both S431 and T432 (ST-KaiC). The concentration of each phosphoform oscillates with a circadian period but with different phases, creating an ordered pattern of phosphoform abundance during each cycle. Multisite phosphorylation of KaiC is required for oscillation, as point mutations at either S431 or T432 abolish rhythmicity (7, 8) (fig. S1).

The predominance of distinct phosphoforms at different points in a cycle—T-KaiC during the phosphorylation phase and S-KaiC during the dephosphorylation phase—suggests that the phosphoform distribution (or a conformation tightly linked to the phosphorylation state) may determine the phase of the oscillator. If this is true, it should be possible to specify the initial phase by preparing KaiC with the appropriate phosphoform distribution

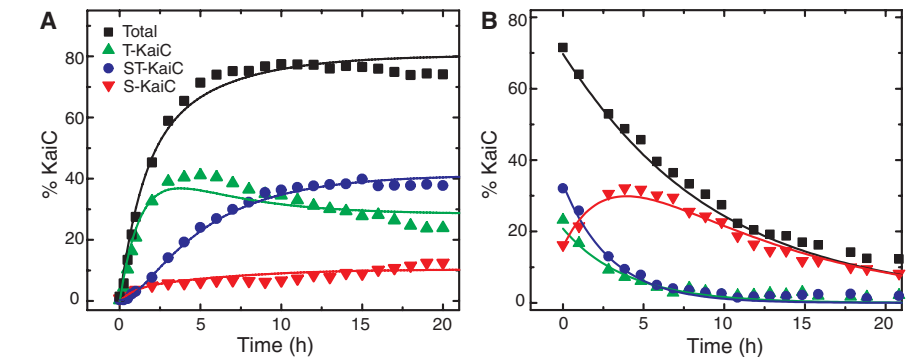


Fig. 2. KaiC phosphoform kinetics in partial reactions. **(A)** KaiC phosphorylation in the presence of KaiA. A least-squares fit (solid lines) to a four-state linear model (Fig. 2C) is shown. **(B)** Autonomous dephosphorylation of KaiC. Phosphorylated KaiC was prepared by incubation with KaiA, which was then removed by immunoprecipitation (28), initiating dephosphorylation. A least-squares fit (solid lines) to the same four-state model, with phosphorylation disallowed, is shown. **(C)** Reaction diagram for the four-state model with first-order kinetics.

and then adding KaiB and KaiA to initiate oscillations. Indeed, a reaction initiated with a KaiC pool enriched in T-KaiC begins in the phosphorylation phase, whereas a reaction initiated with high levels of S-KaiC begins in the dephosphorylation phase (Fig. 1C).

To investigate the origins of the ordered pattern of KaiC phosphoforms, we decomposed the oscillator into partial reactions. When KaiA is mixed with U-KaiC, T-KaiC accumulates first, followed by ST-KaiC, and eventually by S-KaiC (Fig. 2A). When highly phosphorylated KaiC is incubated alone, it autodephosphorylates, and the concentrations of T-KaiC and ST-KaiC decay monotonically (Fig. 2B). Concomitant with the decay of ST-KaiC, the abundance of S-KaiC transiently increases before eventually decaying, which suggests that S-KaiC is produced from ST-KaiC by dephosphorylation of T432. Neither the pattern of dephosphorylation nor its kinetics is

affected by the presence of KaiB (fig. S2). These data are described well by a linear model of KaiC phosphoform interconversion (Fig. 2C). Based on the rate constants calculated by fitting these data (Fig. 2, fig. S3, and table S1), we conclude that KaiA both promotes KaiC phosphorylation (according to a hyperbolic dependence quantified in fig. S4) (11) and inhibits some dephosphorylation steps (6). Phosphoform interconversions occur on the same time scale as the oscillations themselves, consistent with the idea that these are the key slow dynamical processes underlying the oscillator.

Combining the phosphoform dynamics observed in the partial reactions (Fig. 2, A and B) yields the same qualitative pattern of phosphoform abundance observed in the full oscillating reaction, suggesting that each cycle of the oscillator is composed of a phosphorylation phase of high KaiA activity followed by a dephosphorylation phase during which

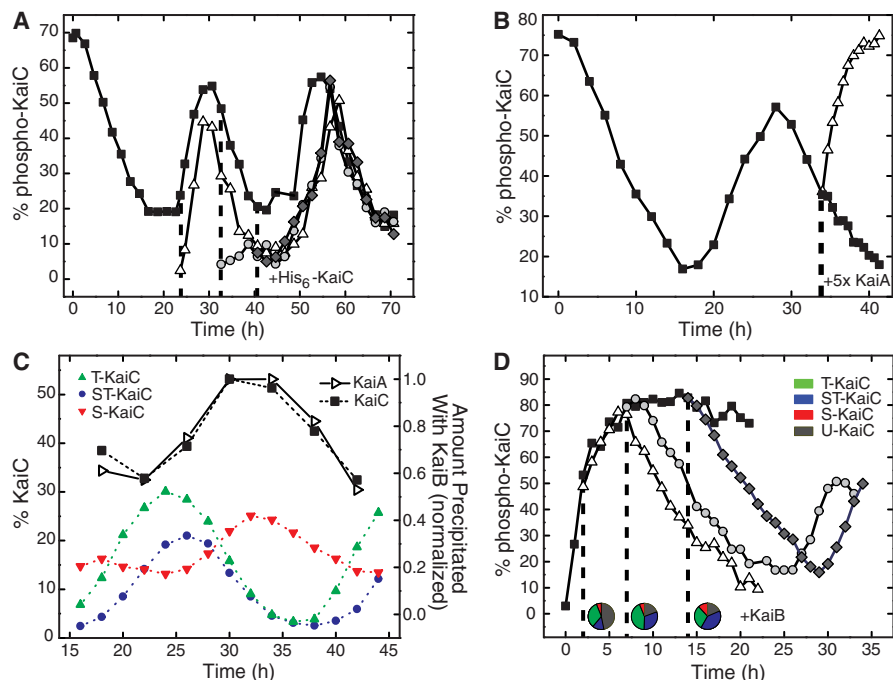


Fig. 3. KaiB suppresses KaiA activity in an S-KaiC dependent manner. **(A)** Global KaiA activity varies during the circadian cycle. Dephosphorylated His₆-KaiC (triangles, circles, diamonds) was added at 10% of the concentration of untagged KaiC (squares) at the indicated times. **(B)** Phosphorylation activity is rapidly restored by the addition of a five-fold excess of KaiA during the dephosphorylation phase (triangles), which continues in a control without excess KaiA (squares). **(C)** KaiB interaction with KaiC and KaiA scales with the abundance of S-KaiC. The right axis shows the normalized amounts of KaiC and KaiA that coimmunoprecipitate with KaiB-FLAG in an oscillating reaction; the left axis shows the corresponding phosphoform distribution. **(D)** KaiB does not affect phosphorylation until S-KaiC is abundant. Dephosphorylated KaiC was incubated with KaiA, and KaiB was introduced into the reaction at various time points (triangles, circles, diamonds) and compared to a control without added KaiB (squares). Pie charts show the KaiC phosphoform distribution at the time of KaiB addition.

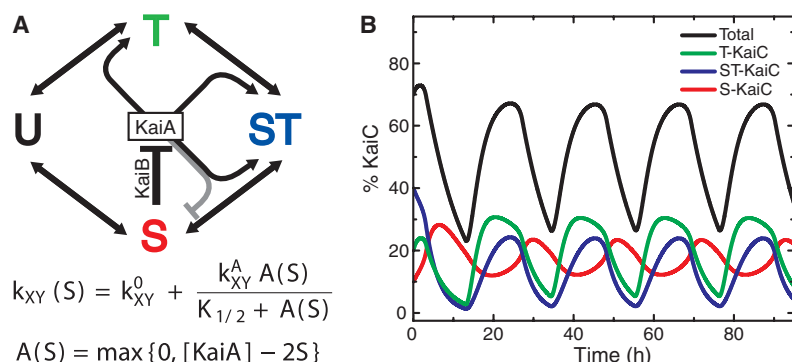


Fig. 4. A model for circadian oscillation driven by multisite KaiC phosphorylation. **(A)** KaiA activity alters the first-order rate constants for interconversion of KaiC phosphoforms. Lines emanating from KaiA ending in an arrowhead (black) or bar (gray) indicate stimulation or repression, respectively, of the transition toward the indicated form of KaiC. We show only the dominant effects (see table S2 and fig. S11). S-KaiC inactivates KaiA via KaiB. The interconversion rates from phosphoform X to Y, k_{XY} , depend hyperbolically on the concentration $A(S)$ of active KaiA monomers, which in turn depends on S-KaiC through its inhibitory activity. See the supporting online text for the complete equations of the dynamical model. **(B)** Numerical integration of the model reproduces circadian oscillation of KaiC phosphorylation.

KaiA is inactive. To test this idea, which has been proposed previously [e.g., (9)], we introduced a small quantity of epitope-tagged, dephosphorylated

KaiC into the oscillator at various times (Fig. 3A); the tagged KaiC will phosphorylate only if KaiA is active. We observe KaiA activity only during the

phosphorylation phase. To rule out the alternative explanation that it is the sensitivity of KaiC to KaiA (21, 23)—rather than the activity of KaiA itself—that varies in a circadian cycle, we increased the concentration of KaiA in the middle of the dephosphorylation phase (Fig. 3B). The phosphorylation level of KaiC immediately increased, indicating that KaiC had not become insensitive to KaiA.

Inactivation of KaiA requires KaiB, as no oscillations occur in its absence (9) (Fig. 2A). Since previous studies have shown that KaiA-KaiB-KaiC complexes form during the clock reaction and that KaiB preferentially binds phosphorylated KaiC (9), we conjectured that inactivation of KaiA and hence initiation of the dephosphorylation phase occurs through a physical interaction between a specific phosphoform of KaiC and KaiA and KaiB. We found that the fraction of KaiC and KaiA complexed with KaiB closely follows the abundance of S-KaiC (Fig. 3C), which suggests that S-KaiC mediates a negative feedback loop through inactivation of KaiA via KaiB. Because S431 and T432 are buried within the KaiC oligomer (25), the specific preference of KaiB for S-KaiC suggests that changes in KaiC phosphorylation may be closely tied to conformational changes sensed by KaiB. To isolate the effects of phosphorylation at each site on KaiB-mediated feedback, we studied single-site nonphosphorylatable KaiC mutants and found that KaiB interacts preferentially with the mutant phosphorylatable only on S431 (fig. S1B) and also preferentially inhibits its phosphorylation (fig. S1A).

To further investigate the timing of KaiB function, we introduced it to a KaiA-KaiC reaction at various points (Fig. 3D). Adding KaiB early in the reaction, when S-KaiC levels are low, does not induce any measurable deviation from the KaiA-KaiC control until S-KaiC has reached ~10% abundance. In contrast, introduction of KaiB when S-KaiC levels are already high (~15%) rapidly induces dephosphorylation. Hence, S-KaiC plays a special role in promoting inactivation of KaiA. Indeed, the similarity between the dephosphorylation pattern caused by adding KaiB when S-KaiC levels are high and that induced by removing KaiA (fig. S5) demonstrates that adding KaiB in the presence of substantial S-KaiC is equivalent to removing KaiA.

To determine whether our understanding of the phosphoform kinetics and feedback mechanism can quantitatively account for the circadian oscillation of KaiC phosphorylation, we created a simple mathematical model (Fig. 4A and fig. S6A) constrained by our experimental measurements. The key assumptions of this minimal model are (i) the concentrations of the three phosphorylated species are the only slow dynamical variables; (ii) the interconversions between phosphoforms are first-order reactions with rates (table S2) that depend hyperbolically on the concentration of active KaiA (Fig. 4A and fig. S4); and (iii) each S-KaiC monomer (together with KaiB) inactivates one KaiA dimer. The phosphorylation and dephosphorylation rates are thus nonlinear functions of the concentration of

S-KaiC—the source of the crucial nonlinear feedback (Fig. 4A).

Using rate constants and a KaiA concentration dependence (table S2) derived solely from data on the non-oscillatory partial reactions (Fig. 2 and figs. S3 and S4), this simple model predicts (Fig. 4B) essential features of the circadian oscillator—period (~21 hours in the model), amplitude of total phosphorylation, sequential appearance of the phosphoforms, and the larger magnitude of the T-KaiC peak (see also fig. S7). This predictive ability suggests that the model captures the key elements of the *in vitro* oscillator. Modifying the model to explicitly treat the formation of KaiA-KaiC complexes (9) likely responsible for promoting KaiC phosphorylation makes it consistent with the observation that the oscillations are rather insensitive to the total concentration of Kai proteins (9) (supporting online text and fig. S8).

The following picture of the origin of stable oscillations emerges (fig. S6A). Starting from the unphosphorylated state, KaiA promotes phosphorylation that is kinetically favored at T432; subsequent phosphorylation at S431 produces ST-KaiC. ST-KaiC can decay via dephosphorylation of T432 to produce S-KaiC, but S-KaiC accumulation is slow because KaiA both inhibits that dephosphorylation and promotes rephosphorylation of S-KaiC to ST-KaiC. Thus, S-KaiC levels remain low until a substantial pool of ST-KaiC has formed. When S-KaiC levels do rise, KaiA activity is reduced, promoting dephosphorylation of ST-KaiC and thereby causing it to rapidly decay into S-KaiC. Thus, S-KaiC accelerates its own production (from ST-KaiC), which causes its concentration to overshoot the point at which KaiA is completely inactivated; this overshoot yields a reservoir of S-KaiC that permits extended inactivation of KaiA even as S-KaiC concentrations decrease through dephosphorylation. In the absence of KaiA activity, T-KaiC and ST-KaiC both dephosphorylate, and S-KaiC—which dephosphorylates more slowly—becomes the dominant remaining phosphorylated species. Eventually enough S-KaiC dephosphorylates for KaiA activity to return, and the cycle begins anew.

To focus on the essential slow dynamics and to be able to derive model parameters directly from our experimental data, our model ignores some known biochemical properties of the Kai proteins and abstracts others into the rate constants. KaiC exists as a hexamer (4), and we have neglected possible effects that depend on the state of the entire hexamer. Further, monomer exchange between hexamers (9) is not explicitly included, and we assume that inhibition of KaiA via KaiB occurs instantaneously upon formation of S-KaiC. In actuality, inhibition appears to take approximately 1 hour (fig. S9), possibly due to slow interaction between KaiB and KaiC or slow exchange of monomers between hexamers. These neglected effects have the potential to increase both the tendency of the system to oscillate and the amplitude of oscillation, but the success of our simplified model suggests that they are not part of the fundamental mechanism.

A recent report from the Kondo group (26) describes the differential phosphorylation of S431 and T432 during the circadian cycle and the interaction of KaiB with KaiC phosphorylated on S431. By using phosphomimetic KaiC mutants, they provide information about ordered phosphorylation complementary to and consistent with our kinetic study of wild-type KaiC.

The most striking behavior of the cyanobacterial circadian oscillator *in vivo* is its precision: Even with asynchronous cell division and an absence of external cues, the clock of a single cell and its offspring maintains precision to a small fraction of a day over several weeks (27). A reductive understanding of the various aspects of the clock—especially that of the core Kai oscillator presented here—should enable us to understand the effects of random fluctuations and variable environments. The *Synechococcus* clock provides an ideal model system for understanding how cells perform quantitative functions in highly variable intra- and extracellular environments.

References and Notes

1. S. L. Harmer, S. Panda, S. A. Kay, *Annu. Rev. Cell Dev. Biol.* **17**, 215 (2001).
2. J. Tomita, M. Nakajima, T. Kondo, H. Iwasaki, *Science* **307**, 251 (2005).
3. M. Nakajima *et al.*, *Science* **308**, 414 (2005).
4. T. Mori *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 17203 (2002).
5. T. Nishiwaki, H. Iwasaki, M. Ishiura, T. Kondo, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 495 (2000).
6. Y. Xu, T. Mori, C. H. Johnson, *EMBO J.* **22**, 2117 (2003).
7. T. Nishiwaki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13927 (2004).
8. Y. Xu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13933 (2004).
9. H. Kageyama *et al.*, *Mol. Cell* **23**, 161 (2006).
10. S. Ye, I. Vakonakis, T. R. Ioerger, A. C. LiWang, J. C. Sacchettini, *J. Biol. Chem.* **279**, 20511 (2004).
11. H. Iwasaki, T. Nishiwaki, Y. Kitayama, M. Nakajima, T. Kondo, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15788 (2002).
12. S. B. Williams, I. Vakonakis, S. S. Golden, A. C. LiWang, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15357 (2002).
13. Y. Kitayama, H. Iwasaki, T. Nishiwaki, T. Kondo, *EMBO J.* **22**, 2127 (2003).
14. H. Iwasaki, Y. Taniguchi, M. Ishiura, T. Kondo, *EMBO J.* **18**, 1137 (1999).
15. H. Kageyama, T. Kondo, H. Iwasaki, *J. Biol. Chem.* **278**, 2388 (2003).
16. S. Clodong *et al.*, *Mol. Syst. Biol.* **3**, 90 (2007).
17. E. Emberly, N. S. Wingreen, *Phys. Rev. Lett.* **96**, 038303 (2006).
18. G. Kurosawa, K. Aihara, Y. Iwasa, *Biophys. J.* **91**, 2015 (2006).
19. A. Mehra *et al.*, *PLoS Comput. Biol.* **2**, e96 (2006).
20. F. Miyoshi, Y. Nakayama, K. Kaizu, H. Iwasaki, M. Tomita, *J. Biol. Rhythms* **22**, 69 (2007).
21. T. Mori *et al.*, *PLoS Biol.* **5**, e93 (2007).
22. H. Takigawa-Imamura, A. Mochizuki, *J. Biol. Rhythms* **21**, 405 (2006).
23. J. S. van Zon, D. K. Lubensky, P. R. Altena, P. R. ten Wolde, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7420 (2007).
24. M. Yoda, K. Eguchi, T. P. Terada, M. Sasai, *PLoS ONE* **2**, e408 (2007).
25. R. Pattanayek *et al.*, *Mol. Cell* **15**, 375 (2004).
26. T. Nishiwaki *et al.*, *EMBO J.* **26**, 4029 (2007).
27. I. Mihalcescu, W. Hsing, S. Leibler, *Nature* **430**, 81 (2004).

Materials and methods are available as supporting material on Science Online.
We thank B. Budnik, J. Neveu, and R. Robinson for assistance with mass spectrometry; T. Mori and C. Johnson for SDS-PAGE conditions for phosphoform separation; J. Ferrell for helpful discussions; and B. Stern, R. Losick, M. Ebert, S. Douglas, and T. Schmidt for comments on the manuscript. This work was supported by an NSF Graduate Research Fellowship (J.S.M.), NSF grant DMR-0229243 (D.S.F.), and the Howard Hughes Medical Institute (E.K.O.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1148596/DC1
Materials and Methods
SOM Text
Figs. S1 to S13
Tables S1 and S2
References

31 July 2007; accepted 21 September 2007
Published online 4 October 2007;
10.1126/science.1148596
Include this information when citing this paper

Disentangling Genetic Variation for Resistance and Tolerance to Infectious Diseases in Animals

Lars Råberg,^{1,2*} Derek Sim,^{1†} Andrew F. Read^{1†}

Hosts can in principle employ two different strategies to defend themselves against parasites: resistance and tolerance. Animals typically exhibit considerable genetic variation for resistance (the ability to limit parasite burden). However, little is known about whether animals can evolve tolerance (the ability to limit the damage caused by a given parasite burden). Using rodent malaria in laboratory mice as a model system and the statistical framework developed by plant-pathogen biologists, we demonstrated genetic variation for tolerance, as measured by the extent to which anemia and weight loss increased with increasing parasite burden. Moreover, resistance and tolerance were negatively genetically correlated. These results mean that animals, like plants, can evolve two conceptually different types of defense, a finding that has important implications for the understanding of the epidemiology and evolution of infectious diseases.

Defense against pathogenic microorganisms and other parasites can be divided into two conceptually different components: resistance (the ability to limit parasite burden) and toler-

ance (the ability to limit the disease severity induced by a given parasite burden) (1–4). It is important to distinguish between these two components because, by definition, resistance has a negative effect on

parasites, whereas tolerance does not; as a result, their relative importance will have substantial consequences for the ecology and evolution of host-parasite interactions (2, 5–7). The distinction between resistance and tolerance has attracted considerable attention in studies of the evolution of plant defense against parasites and herbivores (3, 6, 8). In this area, the emerging pattern is that plants generally exhibit genetic variation for both resistance and tolerance (1, 6, 9, 10).

When it comes to animals, numerous studies have demonstrated genetic variation for resistance, where resistance is typically measured as the inverse of parasite burden (11–14). However, little is known about whether animals may also show genetic variation for tolerance. Yet together, resistance and tolerance are the two components of antipathogen defense that determine disease severity. Suggestive evidence for tolerance in animals comes from a study of α^+ -thalassemia, a monogenic hemoglobin disorder in humans that protects against malaria. Individuals that are hetero- or homozygous for this mutation do not have lower infection intensities of *Plasmodium falciparum* than individuals that are homozygous for the wild type, but the degree of anemia at high infection intensities is diminished, thereby reducing mortality from malaria (15). Thus, it seems α^+ -thalassemia affects tolerance but not resistance to *P. falciparum*. So far as we are aware, no study has yet formally disentangled genetic variation in these two components of defense in any animal host-parasite system.

In the plant literature, tolerance is usually defined as the slope of host fitness against infection intensity (1, 16, 17). In other words, the tolerance of a host genotype is its reaction norm to infection intensity. A tolerant genotype is one in which disease severity is relatively unaffected by increasing pathogen burden, whereas the fitness of a less tolerant genotype declines more rapidly as pathogen burdens rise (Fig. 1). If the reaction norms of different host genotypes vary (that is, if there is a statistical interaction between host genotype and infection intensity), then there is genetic variation for tolerance. We have borrowed this approach to defining and measuring genetic variation in tolerance from the plant literature and applied it to a malaria model system (*Plasmodium chabaudi* in laboratory mice) to investigate whether animal hosts may show genetic variation for tolerance and whether resistance and tolerance are correlated traits.

P. chabaudi is widely used as a model of human malaria (18, 19). Previous studies have shown that there is considerable variation among mouse strains (i.e., genetic variation) for resistance to *P. chabaudi* (20–22). To investigate whether there is also genetic

variation for tolerance, we performed an experiment with five different inbred mouse strains (23). Mice were infected with one of three different *P. chabaudi* clones or left uninfected in a fully factorial design. The experiment was performed in three experimental blocks separated in time. As with human malaria, one of the main causes of morbidity and mortality in rodent malaria is anemia. *P. chabaudi* also causes weight loss in mice. The degree of red blood cell (RBC) loss and weight loss is correlated with infection intensity and predicts mortality (24). To test for variation in tolerance, we therefore used minimum RBC density and minimum weight during the infection as measures of disease severity (analogous to host fitness used in the plant literature). Specifically, we tested whether the slopes of the relations be-

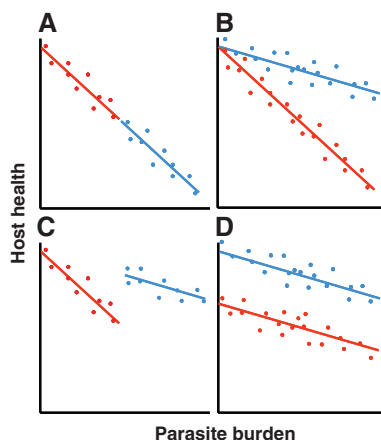


Fig. 1. Schematic figure showing reaction norms of two host genotypes (red or blue line) for disease severity across a range of infection intensities in individual hosts (dots). (A) Two equally tolerant genotypes differing in resistance; here, the red genotype has lower parasite burdens (is more resistant) and thereby maintains a higher health status when infected. (B) Two equally resistant genotypes (same average parasite burden), but here the red genotype is less tolerant (health declines faster with increasing parasite burden). (C) Genotypes differ in both tolerance and resistance; here, the more tolerant genotype (blue) is less resistant, so that both genotypes end up having, on average, the same health status when infected. (D) Host genotypes differ in neither resistance (same average parasite burden) nor tolerance (same slope). Instead, the genetic difference in health status is due to a difference in intercept, so that the difference exists even when animals are uninfected. It is thus indicative of genetic differences in “general vigor” (8) and has nothing to do with defense against the infectious agent in question. Because of the possible existence of variation in general vigor, tolerance has to be defined as a reaction norm, and so it can only be measured and compared across groups of animals (17). Thus, in contrast to resistance, it is not possible to compare the tolerance of two individual hosts. Demonstrating genetic variation for tolerance therefore requires that disease severity be assessed in animals of the same genotype across a range of infection intensities; a difference in slope between genotypes indicates genetic variation for tolerance.

tween infection intensity and minimum RBC density or minimum weight differed between mouse strains.

As usual in this host-parasite system (25, 26), there was a distinct peak in parasite density around day 8 post-inoculation. Minimum RBC density occurred

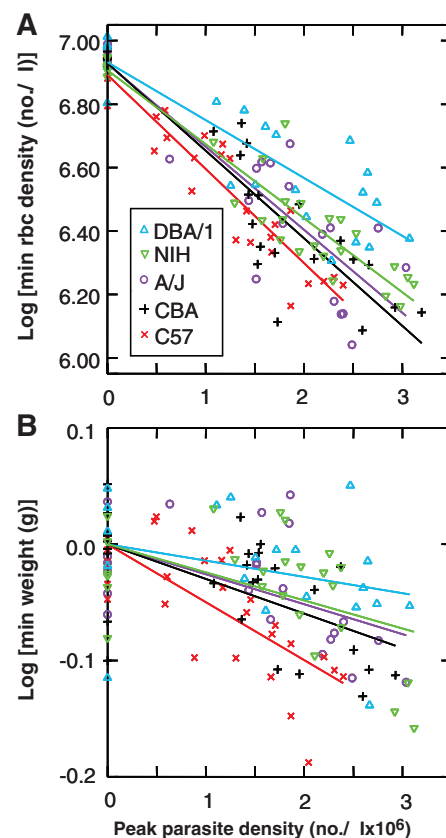


Fig. 2. Variation for tolerance among mouse strains. (A) Minimum RBC density (log-transformed) versus peak parasite density. Mouse strain \times parasite density: $F_{4,117} = 6.08$, $P = 0.0002$; parasite density: $F_{1,117} = 173.3$, $P < 0.0001$; mouse strain: $F_{4,117} = 0.20$, $P = 0.94$; and experimental block: $\chi^2 = 22.1$, $P < 0.0001$. Initial RBC density [$F_{1,116} = 0.80$, $P = 0.37$], the quadratic terms [parasite density: $F_{1,117} = 0.76$, $P = 0.38$; strain \times parasite density: $F_{4,111} = 0.33$, $P = 0.86$], and the interaction between block and strain ($P > 0.25$) were not significant and were therefore excluded from the model. (B) Minimum weight (log-transformed) versus peak parasite density. Strain \times parasite density: $F_{4,110} = 6.06$, $P = 0.0002$; parasite density: $F_{1,111} = 8.09$, $P = 0.0053$; parasite density: $F_{1,111} = 34.4$, $P < 0.0001$; mouse strain: $F_{4,110} = 2.76$, $P = 0.031$; initial weight: $F_{1,111} = 140$, $P < 0.0001$; and experimental block: $\chi^2 = 18.1$, $P < 0.0001$. Strain \times parasite density: $F_{4,105} = 1.20$, $P = 0.31$ and the interaction between block and strain ($P > 0.25$) were not significant, and these terms were therefore excluded. To facilitate the comparison of slopes and because the initial weight (the intercept) differed between strains but for the present purposes is an irrelevant main effect when testing for resistance and tolerance, the reaction norms for weight have been scaled so that all genotypes have an intercept of zero. DBA/1, NIH, A/J, CBA, and C57 are the different mouse strains.

¹Institute of Evolutionary Biology and Institute of Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK. ²Department of Animal Ecology, Lund University, Ecology Building, 223 62 Lund, Sweden.

*To whom correspondence should be addressed. E-mail: lars.raberg@zoekol.lu.se

†Present address: Center for Infectious Disease Dynamics, Departments of Biology and Entomology, Pennsylvania State University, 208 Mueller Lab, University Park, PA 16802, USA.

around day 11, on average 2.49 ± 0.88 (mean \pm SD) days after the peak parasite density. Minimum weight occurred around days 10 to 11, on average 1.60 ± 2.89 days after the peak parasite density (fig. S1).

To test for variation in resistance among mouse strains, we performed an analysis of peak parasite density against mouse strain and parasite clone. Peak parasite density differed between mouse strains [$F_{4,102} = 15.54$, $P < 0.0001$] and parasite clones [$F_{2,103} = 64.81$, $P < 0.0001$], but there was no strain-by-clone interaction [$F_{8,102} = 0.66$, $P = 0.73$]. There was also a significant effect of experimental block ($\chi^2 = 47.4$, $P < 0.0001$), but no interactions between block and strain and/or clone ($P > 0.25$). Thus, as in previous studies (20–22, 25), mouse strains differed in resistance, and parasite clones differed in the infection intensity that they induced.

To test for genetic variation for tolerance, we performed analyses of minimum RBC density and minimum weight against peak parasite density (both linear and quadratic terms), mouse strain, their interactions, and pre-inoculation values of RBC density or weight. In the case of both RBC and weight loss, there were highly significant interactions between strain and parasite density (Fig. 2). Thus, there was variation among mouse strains in tolerance measured in terms of either anemia or weight loss. This conclusion is robust to the inclusion of parasite clones in the statistical models, the exclusion of

uninfected animals from the analyses, or the use of different infection intensity measures (see supporting online material text).

To test whether these two estimates of tolerance were correlated, we calculated the slopes of minimum weight and RBC density against the peak parasite density for each mouse strain. There was a significant correlation between the two measures of tolerance (Spearman's rank correlation, $r_s = 1.0$, $n = 5$ mouse strains, $P < 0.05$). There were also significant negative correlations between resistance and both measures of tolerance ($r_s = -1.0$, $n = 5$, $P < 0.05$ in the case of both RBC loss and weight loss) (Fig. 3). DBA mice, for example, were more tolerant and less resistant than C57s, which were the opposite. Thus, reduced tolerance is a cost of resistance and vice versa.

Our studies demonstrate that the conceptual and analytical framework developed by plant evolutionary biologists can also be used to reveal genetic variation for tolerance to infectious diseases in animals. The existence of genetic variation for both resistance and tolerance means that host defense can take a variety of evolutionary trajectories in response to pathogen pressure. The mechanistic basis of the genetic variation in tolerance we report remains to be determined. Variation in tolerance measured as RBC loss could occur because either the rate of regeneration of RBCs or the rate of destruction of RBCs by parasites and/or host immune responses varies among strains. The correlation between tolerance measured as RBC and weight loss suggests that there is a common underlying factor between these two forms of tolerance. In plants, where tolerance has long been studied, genes conferring disease tolerance have yet to be identified at the molecular level (6). In our disease model, resistance and tolerance were traded off against each other (Fig. 3). A similar trade-off has previously been demonstrated in the context of plant defense against herbivory (2). In the case of infectious diseases, a trade-off could arise if the price of more aggressive immune control of infection is increasing collateral damage (immunopathology).

Our findings, if they prove to be general, have important implications for our understanding of the ecology and evolution of animal host-parasite interactions. First, whereas the evolution of resistance has a negative effect on the prevalence of the infectious agent in the host population, tolerance should have a neutral or positive effect. Thus, resistance and tolerance have contrasting effects on the epidemiology of infectious diseases (5, 7). Second, hosts and parasites are commonly thought to be engaged in antagonistic coevolution, where evolution of host resistance selects for counteradaptations in the parasite, which selects for improved resistance in the host and so on, leading to open-ended nonequilibrium evolutionary dynamics (27). However, tolerance does not have a negative effect on the fitness of the parasite, and so it cannot fuel antagonistic coevolution in the same way as is expected of resistance. Genetic variation for tolerance can therefore be expected to allow the sort of host evolution that will substantially dampen antagonistic coevolution (6).

Beyond evolutionary ecology, there is a clear need to recognize and separate the two components

of disease defense in the context of animal breeding. For instance, attempts to enhance yield in agricultural animals by artificial selection on disease resistance traits or on total yield in the face of infection [often referred to as “resilience” (28)] could generate a variety of more or less desirable outcomes, depending on how resistance or yield varies with tolerance. The experimental and analytic approach used here is readily transferable to domestic animals where it could be used to work out optimal selection strategies.

References and Notes

1. E. L. Simms, J. Triplett, *Evol. Int. J. Org. Evol.* **48**, 1973 (1994).
2. W. Fineblum, M. Rausher, *Nature* **377**, 517 (1995).
3. J. Fornoni, J. Nuñez-Farfan, P. L. Valverde, M. Rausher, *Evol. Int. J. Org. Evol.* **58**, 1685 (2004).
4. O. Restif, J. C. Koella, *Am. Nat.* **164**, E90 (2004).
5. B. A. Roy, J. W. Kirchner, *Evol. Int. J. Org. Evol.* **54**, 51 (2000).
6. M. D. Rausher, *Nature* **411**, 857 (2001).
7. M. R. Miller, A. White, M. Boots, *Evol. Int. J. Org. Evol.* **60**, 945 (2006).
8. K. Stowe, R. Marquis, C. Hochwender, E. L. Simms, *Annu. Rev. Ecol. Syst.* **31**, 565 (2000).
9. P. X. Kover, B. A. Schaal, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11270 (2002).
10. T. Koskela, S. Puustinen, V. Salonen, P. Mutikainen, *Evol. Int. J. Org. Evol.* **56**, 899 (2002).
11. D. Malo, E. Skamene, *Trends Genet.* **10**, 365 (1994).
12. A. V. S. Hill, *Annu. Rev. Immunol.* **16**, 593 (1998).
13. J. A. Smith, K. Wilson, J. G. Pilkington, J. M. Pemberton, *Proc. R. Soc. London Ser. B Biol. Sci.* **266**, 1283 (1999).
14. T. N. Williams *et al.*, *Nat. Genet.* **37**, 1253 (2005).
15. S. Wambua *et al.*, *PLoS Medicine* **3**, e158 (2006).
16. P. Tiffin, M. D. Rausher, *Am. Nat.* **154**, 700 (1999).
17. E. L. Simms, *Evol. Ecol.* **14**, 563 (2000).
18. M. M. Stevenson, E. M. Riley, *Nat. Rev. Immunol.* **4**, 169 (2004).
19. A. Martinelli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 814 (2005).
20. M. M. Stevenson, J. J. Lyanga, E. Skamene, *Infect. Immun.* **38**, 80 (1982).
21. A. Fortin, M. M. Stevenson, P. Gros, *Genes Immun.* **3**, 177 (2002).
22. K. Grech, K. Watt, A. F. Read, *J. Evol. Biol.* **19**, 1620 (2006).
23. Materials and methods are available as supporting material on Science Online.
24. M. J. Mackinnon, A. F. Read, *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **359**, 965 (2004).
25. M. J. Mackinnon, A. F. Read, *Evol. Int. J. Org. Evol.* **53**, 689 (1999).
26. J. C. De Rooode *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7624 (2005).
27. M. E. J. Woolhouse, J. P. Webster, E. Domingo, B. Charlesworth, B. Levin, *Nat. Genet.* **32**, 569 (2002).
28. S. A. Bisset, C. A. Morris, *Int. J. Parasitol.* **26**, 857 (1996).
29. We thank A. Graham, H. Smith, and M. Stjernman for extensive discussion. This work was funded by the Wellcome Trust (A.F.R.), a Marie Curie fellowship (FP6-501567) (L.R.), and the Swedish Research Council (L.R.). This paper was written while A.R. was at the Wissenschaftskolleg zu Berlin. The experimental work was conducted on Project License 60/2714 granted by the Home Office under the auspices of the U. K. Animals (Scientific Procedures) Act of 1986.

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/812/DC1

Material and Methods

SOM Text

Fig. S1

References

30 July 2007; accepted 3 October 2007

10.1126/science.1148526

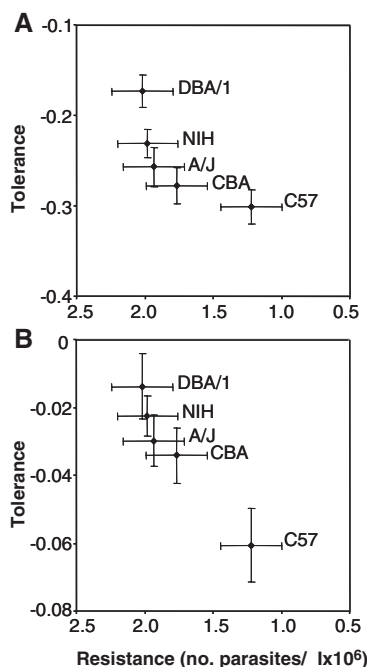


Fig. 3. Trade-off between tolerance and resistance. (A) Correlation between resistance (inverse of peak parasite density) and tolerance in the form of minimum RBC density during infection (that is, tolerance measured as the slope of a regression of minimum RBC density against peak parasite density). (B) Correlation between resistance and tolerance in the form of minimum weight during infection (that is, tolerance measured as the slope of a regression of minimum weight against peak parasite density). Plots show mean \pm SEM for each mouse strain.

TARP Auxiliary Subunits Switch AMPA Receptor Antagonists into Partial Agonists

Karen Menuz,^{1,2} Robert M. Stroud,^{3*} Roger A. Nicoll,^{1,4*} Franklin A. Hays³

Quinoxalinedione compounds such as 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX) are the most commonly used α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor antagonists. However, we find that in the presence of transmembrane AMPA receptor regulatory proteins (TARPs), which are AMPA receptor auxiliary subunits, CNQX acts as a partial agonist. CNQX induced small depolarizing currents in neurons of the central nervous system, and reconstitution of this agonist activity required coexpression of TARPs. A crystal structure of CNQX bound to the TARP-less AMPA receptor ligand-binding domain showed that, although CNQX induces partial domain closure, this movement is not transduced into linker separation, suggesting that TARPs may increase agonist efficacy by strengthening the coupling between domain closure and channel opening. Our results demonstrate that the presence of an auxiliary subunit can determine whether a compound functions as an agonist or antagonist.

Excitatory synaptic transmission in the brain is mediated by glutamate acting on two classes of ionotropic receptors: AMPA and *N*-methyl-D-aspartate (NMDA) receptors. A major breakthrough in the field of excitatory synaptic transmission came with the discovery of the quinoxalinedione series of competitive AMPA receptor antagonists in 1988 (1–3). These drugs—CNQX, 6,7-dinitroquinoxaline-2,3-dione (DNQX), and 2,3-dioxo-6-nitro-1,2,3,4-tetrahydrobenzo[*f*]quinoxaline-

7-sulfonamide (NBQX)—potently and selectively block AMPA receptors and have been indispensable in characterizing excitatory synaptic transmission in the central nervous system (CNS).

As competitive antagonists, quinoxalinediones interact with AMPA receptors in the same binding pocket as agonists, thereby occluding agonist binding. Unlike agonists, competitive antagonists have no efficacy, meaning that binding does not lead to opening of the ion channel (i.e., gating). Despite the current widespread use of quinoxalinediones as AMPA receptor antagonists to study synaptic transmission, a few reports indicate that these compounds can have excitatory actions on a subset of interneuron populations, though not on excitatory neurons (4–7). The mechanism for this effect remains unexplained.

While recording cerebellar granule cells, we also observed that application of CNQX (10 μ M) increased spontaneous inhibitory postsynaptic current (sIPSC) frequency (0.22 \pm 0.07 Hz to 1.58 \pm 0.56 Hz,

$n = 5$ cells, $P < 0.05$) without a change in amplitude (21.1 \pm 6.1 pA to 19.3 \pm 3.1 pA, $n = 5$ cells) (Fig. 1, A and B) (5), implying increased excitability of the presynaptic interneurons: cerebellar Golgi cells (8). Application of CNQX to Golgi cells reliably evoked an inward current in voltage-clamp recordings (–37.9 \pm 4.4 pA, $n = 11$ cells), indicating a non-canonical depolarizing effect of CNQX on these cells (Fig. 1C). However, we found that antagonizing AMPA receptors does not simply lead to depolarization because application of GYKI 53655 (10 μ M), which is a noncompetitive AMPA receptor antagonist (9), did not induce an inward current in Golgi cells (fig. S1). Given that the CNQX-induced depolarization was recorded in γ -aminobutyric acid type A (GABA_A), NMDA, and glycine receptor antagonists as well as in tetrodotoxin (TTX) to block network activity, our data suggested that CNQX was not simply acting through another neuronal receptor but instead that CNQX may act as an AMPA receptor agonist on these cells.

We therefore tested whether a noncompetitive AMPA receptor antagonist could block the CNQX-induced depolarization and whether a positive allosteric modulator could potentiate the response. Preincubation of cerebellar slices with GYKI 53655 blocked the CNQX-induced current (CNQX-induced current in GYKI = 0.4 \pm 2.5 pA, $n = 5$ cells, $P < 0.001$, as compared with CNQX alone) (Fig. 1C). Furthermore, trichloromethiazide (TCM) (500 μ M), a positive modulator structurally similar to cyclothiazide (10, 11), increased the response to CNQX (–88.7 \pm 21.6 pA, $n = 5$ cells, $P < 0.01$, as compared with CNQX alone) (Fig. 1D). Thus, CNQX appears to act as an agonist on these AMPA receptors, despite its previous characterization as a competitive antagonist.

Although AMPA receptors are expressed on most, if not all, neurons in the brain, a depolarizing action of CNQX was not previously reported for

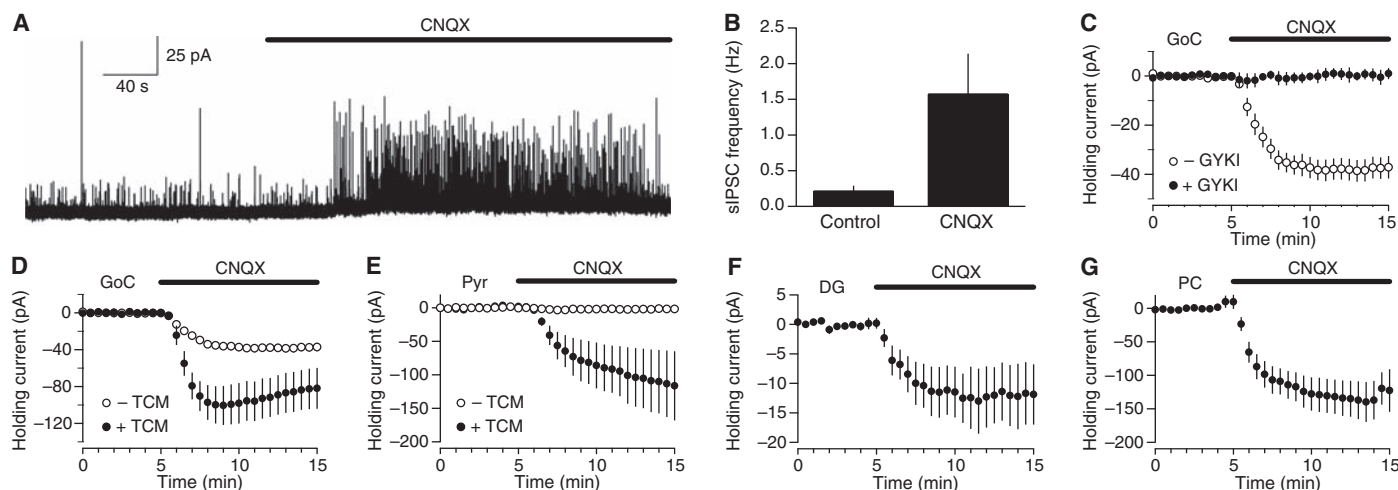


Fig. 1. Depolarizing current elicited by CNQX in neurons. (A and B) The frequency of sIPSCs in cerebellar granule cells voltage-clamped at 0 mV [excitatory postsynaptic current (EPSC) reversal potential] was measured before and after bath application of CNQX ($n = 5$ cells, $P < 0.05$). (C and D) The holding current needed to voltage-clamp cerebellar Golgi cells (GoC) to –70 mV was measured as CNQX was applied in the absence or presence of either 10 μ M GYKI 53655 ($n = 11$ and 5 cells, respectively, $P < 0.001$) (C) or 500 μ M TCM ($n = 11$

and 5 cells, respectively, $P < 0.01$) (D). (E) Similarly, CNQX was applied to hippocampal CA1 pyramidal cells (Pyr) in the presence and absence of TCM ($n = 5$ cells for each treatment). (F and G) CNQX was also applied to hippocampal dentate granule cells (DG) ($n = 5$ cells) and Purkinje cells (PC) ($n = 5$ cells) in the presence of TCM. Hippocampal neurons were voltage-clamped at –70 mV; Purkinje neurons were held at –10 mV to prevent voltage escape. Error bars in (B) to (G) indicate SEM.

excitatory neurons (5, 6). This raised the possibility that AMPA receptors on interneurons somehow differ from those on other types of neurons. We therefore tested whether CNQX can act as an agonist on AMPA receptors expressed by hippocampal CA1 pyramidal cells. In the absence of TCM, CNQX did not evoke an inward current (-1.9 ± 2.6 pA, $n = 5$ cells) (Fig. 1E), as was reported previously (6, 12). However, a CNQX-induced current was observed in the presence of TCM (-103.5 ± 43.3 pA, $n = 5$ cells) (Fig. 1E). A CNQX-induced current was also detected in the presence of TCM in dentate granule cells (-12.2 ± 5.1 pA, $n = 5$ cells) (Fig. 1F) and in cerebellar Purkinje cells (-131.7 ± 27.2 pA, $n = 5$ cells) (Fig. 1G). Given that CNQX could induce depolarizing currents in all neuron types tested, the agonist activity of CNQX is most likely a general property of CNS neurons and AMPA receptors. The previously reported lack of detection of CNQX-induced currents was most likely because CNQX was not tested in the presence of TCM.

To conclusively attribute the depolarizing current to AMPA receptor activation, we next attempted to reconstitute the agonist activity of CNQX on AMPA receptors expressed in cultured human embryonic kidney (HEK) 293 cells. A brief application of CNQX evoked an inward current on HEK293 cells transfected with the AMPA receptor pore-forming subunit GluR1(Q) flop and γ -2, a member of the TARP family of AMPA receptor auxiliary subunits (13–15) (-12.7 ± 5.7 pA, $n = 6$ cells) (Fig. 2A). Furthermore, the CNQX-induced current in HEK293 cells had many properties consistent with AMPA receptor activation. The response to CNQX was significantly enhanced in the presence of TCM (-287.0 ± 104.6 pA, $n = 6$ cells, $P < 0.04$) (Fig. 2A). For ease of measurement, all further experiments in HEK293 cells were carried out in the presence of TCM. Comparison of the currents evoked by CNQX and glutamate (1 mM) indicated that CNQX is acting as a partial agonist (Fig. 2A). Partial agonists, such as kainate, bind AMPA receptors but only induce a fraction of the activation induced by full agonists, such as glutamate. As in Golgi cells, GYKI 53655 blocked the CNQX-induced inward current in HEK293 cells (inhibition: $99.3 \pm 0.7\%$, $n = 4$ cells) (Fig. 2B), and the effect of GYKI was reversible. Furthermore, the current elicited by CNQX had the current-voltage (I - V) relationship expected for GluR1(Q) AMPA receptors ($n = 4$ cells) (Fig. 1C) (16). Together, our data indicated that CNQX acts as a partial agonist in a heterologous expression system.

We sought to determine whether other members of the quinoxalinedione family also have agonist activity. A brief application of DNQX elicited an inward current in transfected HEK293 cells (-284.0 ± 99.4 pA, $n = 13$ cells) (Fig. 2D), which was not statistically different from that seen with CNQX (-211.1 ± 69.9 pA, $n = 11$ cells, $P = 0.57$). In contrast, NBQX did not elicit an inward current in HEK293 cells (23.2 ± 7.0 pA, $n = 9$ cells) (Fig. 2D). We observed similar effects in cerebellar Golgi cells in the absence of TCM (DNQX: -10.0 ± 3.0 pA, $n = 6$ cells; NBQX: 2.1 ± 2.1 pA, $n = 5$ cells) (Fig. 2E). Therefore, NBQX acts purely as a competitive AMPA

Fig. 2. AMPA receptor activation by CNQX and DNQX, but not NBQX, in a heterologous system. (A) The change in holding current, elicited by a brief application of 10 μ M CNQX to voltage-clamped (-70 mV) HEK293 cells expressing GluR1 flop and γ -2, was measured in the absence and presence of 500 μ M TCM. A brief application of 1 mM glutamate (Glut) in TCM was later applied ($n = 6$ cells). (B) The CNQX-induced change in holding current was measured before, during, and after the application of 100 μ M GYKI 53655 ($n = 4$ cells). (C) The I - V relationship of the CNQX-induced current in TCM is shown ($n = 4$ cells).

(D and E) The effects of 10 μ M DNQX and 10 μ M NBQX were also measured in HEK293 cells cotransfected with GluR1 and γ -2 (DNQX: $n = 11$ cells; NBQX: $n = 9$ cells) (D) and in Golgi cells in the absence of TCM (DNQX: $n = 6$ cells; NBQX: $n = 5$ cells) (E). Error bars in (A) to (E) indicate SEM.

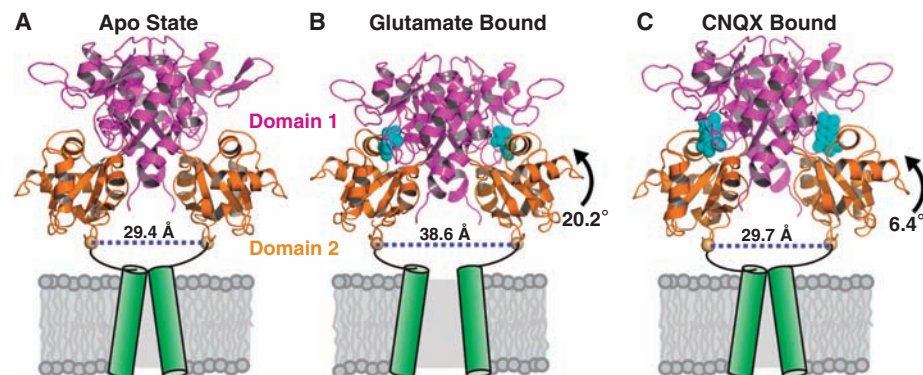
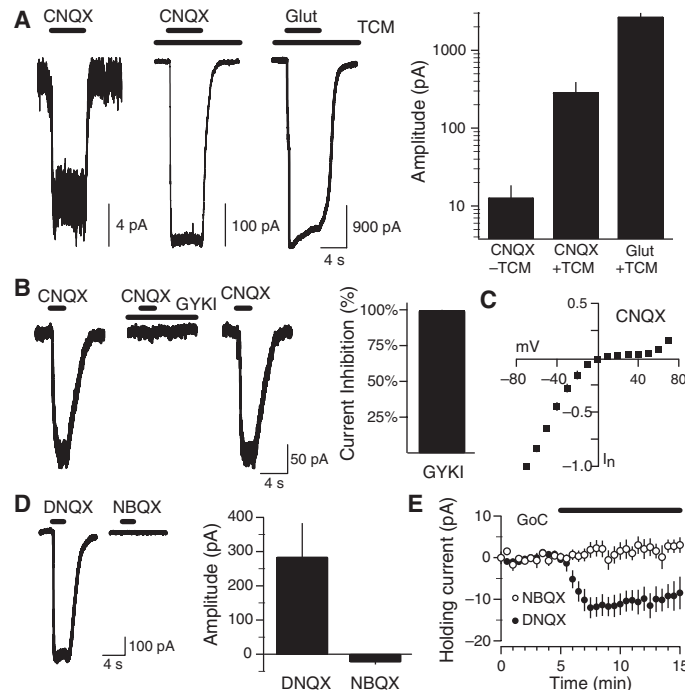


Fig. 3. Crystal structure of CNQX bound to the AMPA receptor LBD. iGluR monomers are composed of three transmembrane helices and one reentrant loop, an N-terminal domain, and a LBD. Shown are two pore-forming helices (rendered as green cylinders) within a membrane bilayer and the LBD [rendered as a ribbon diagram and colored in magenta (domain 1) and orange (domain 2)]. The LBD is a back-to-back dimer of two protomers, each of which contains a domain 1 and domain 2. (A and B) GluR2 LBD crystal structure in the *apo* and glutamate-bound states [Protein Data Bank codes 1FT0 and 1FT1, respectively (32)]. Glutamate is rendered as a Corey-Pauling-Koltun (CPK) model and is colored in cyan in (B). Domain closure denotes the degree of domain 2 movement toward domain 1 relative to the *apo* state (20.2°). Linker distance is measured between the C_α atoms of Pro⁶³² in each protomer. In the full-length receptor, the artificial linker (orange sphere) is connected to the pore-forming helices. Glutamate binding induces domain closure, increases the linker distance, and opens the channel pore. (C) GluR2 LBD crystal structure in the CNQX-bound state (rendered as a CPK model and colored in cyan).

receptor antagonist, whereas both CNQX and DNQX act as partial agonists.

Models of channel activation and desensitization have been developed through structural studies of isolated ionotropic glutamate receptor ligand-binding domains (LBDs). These LBDs consist of two domains arranged in a clamshell-like manner that undergo a conformational rearrangement, typically movement of domain 1 toward domain 2 (domain closure), upon ligand binding (Fig. 3). Binding of full agonists, such as

glutamate, induces maximal domain closure ($\sim 21^\circ$), whereas partial agonists, such as kainate, induce partial closure (12°) relative to the unbound *apo* state (17). Therefore, channel activation is correlated with the degree of domain closure upon ligand binding (17, 18). Agonist efficacy also correlates with the length of separation between the linker regions of each subunit, which connect the LBDs to the pore-forming transmembrane segments in the full-length receptor (Fig. 3) (17–20).

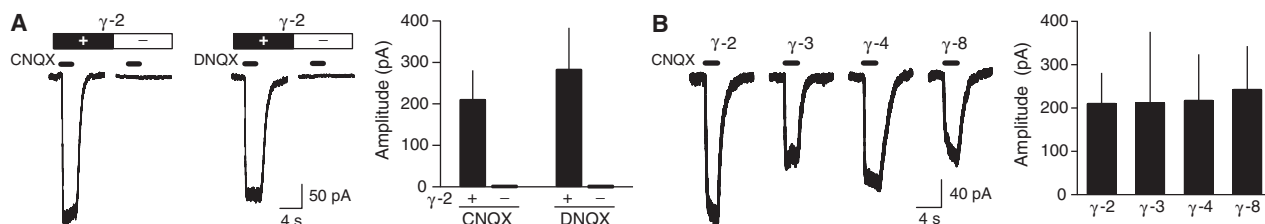


Fig. 4. Effects of CNQX and DNQX in the absence and presence of TARPs. **(A)** Either CNQX or DNQX was applied in the presence of TCM to HEK293 cells cotransfected with GluR1 and γ -2 or with GluR1 alone [CNQX: GluR1 + γ -2 ($n = 11$ cells) and GluR1 alone ($n = 8$ cells); DNQX: GluR1 + γ -2 ($n =$

13 cells) and GluR1 alone ($n = 12$ cells)]. **(B)** CNQX-induced holding current changes in cells transfected GluR1 and either TARP γ -2, γ -3, γ -4, or γ -8 ($n = 11, 4, 5$, and 4 cells, respectively). Error bars in **(A)** and **(B)** indicate SEM.

To test whether CNQX induces domain closure and linker separation consistent with agonist activity, we obtained the crystal structure of CNQX bound to the GluR2 LBD, also referred to as "S1S2" (Fig. 3C, table S1, and fig. S2). The CNQX-bound structure was $\sim 6.4^\circ$ closed relative to the *apo* state ($\sim 7.7^\circ$ and $\sim 5.1^\circ$ for the two protomers), which is consistent with partial agonist activity. However, the linker separation in the CNQX-bound structure (29.7 Å) was not different from that of the *apo* state (29.4 Å), suggesting that CNQX may not transduce its domain closure into channel opening, and is therefore inconsistent with the partial agonist activity that we observed. Our CNQX-bound structure was similar to the structure of DNQX bound to the GluR2 LBD (C α root mean square deviation = 0.71 Å) (18).

A key difference between the CNQX-bound LBD structure and native receptors is that the latter are coexpressed with TARPs, which include γ -2, γ -3, γ -4, and γ -8. To determine whether coexpression of TARPs is required for CNQX-induced receptor activation, we compared the CNQX-induced currents in HEK293 cells transfected with GluR1 and γ -2 to those transfected with GluR1 alone. Neither CNQX nor DNQX activated AMPA receptors in the absence of γ -2, even in the presence of TCM (CNQX: -1.63 ± 1.10 pA, $n = 8$ cells; DNQX: -1.08 ± 0.56 pA, $n = 12$ cells) (Fig. 4A). This was not due to a lack of AMPA receptor expression because the average glutamate-evoked currents in cells transfected with GluR1 alone were similar in the two conditions (GluR1 alone: 2941 ± 485 pA, $n = 20$ cells; γ -2 + GluR1: 3386 ± 409 pA, $n = 24$ cells; $P = 0.48$). Furthermore, the use of TCM to both slow and block AMPA receptor desensitization (10, 11) suggests that the lack of CNQX-induced currents in the absence of γ -2 is not simply due to a detection difficulty resulting from rapid desensitization. This indicates that CNQX and DNQX act as pure competitive antagonists on AMPA receptors in the absence of γ -2 but as agonists in the presence of γ -2. Other members of the TARP family produced similar changes in CNQX activity (Fig. 4B). Therefore, inclusion of any TARP family member in an AMPA receptor complex switches the nature of CNQX and DNQX from competitive antagonists to partial agonists.

In light of our results, the structurally similar quinoxalinediones—CNQX and DNQX—are perhaps best considered to be partial agonists rather than competitive antagonists, given that most neurons in the CNS express TARPs (21, 22). The lack of CNQX agonist activity in the absence of TCM in

hippocampal CA1 pyramidal cells (Fig. 1E) (6) and other excitatory neurons may be due to rapid desensitization of CNQX-induced currents in these cells as a result of the expression of the flip splice variant of AMPA receptor subunits (23, 24). The current induced by CNQX on GluR1 flip receptors was only $\sim 3\%$ of the current induced by CNQX in the presence of TCM (Fig. 2A and fig. S3B). If hippocampal CA1 pyramidal cells primarily express flip receptors, we would expect only 3 pA of CNQX-induced current in the absence of TCM, which is similar to the trend that we observed (-1.9 ± 2.6 pA). In contrast, the weaker TCM potentiation in Golgi cells may be due to expression of the flop variants of AMPA receptors, because the CNQX-induced currents from GluR1 flop receptors were only moderately potentiated by TCM (fig. S3). This moderate potentiation agrees with the known weaker potency of cyclothiazide, a congener of TCM, on flop receptors (25).

Although CNQX acts as a weak partial agonist, its occupancy of the AMPA receptor ligand-binding site would prevent the binding of other agonists and further activation. Thus, interpretations of previous experiments that used CNQX to antagonize AMPA receptors are valid, except perhaps in the case when neurons were incubated for extended periods with CNQX, which may lead to cell death and/or receptor internalization resulting from chronic AMPA receptor activity. Given that quinoxalinedione analogs, as well as other competitive antagonists of AMPA receptors, have undergone clinical testing for the treatment of numerous diseases, including epilepsy and stroke (26–28), better predictive activity of *in vivo* function may benefit from the inclusion of TARPs in *in vitro* drug screening assays.

Although previous work had shown that TARPs increase the efficacy of AMPA receptor agonists (29–31), their relationship to our structure- and function-based understanding of AMPA receptor gating had remained entirely unknown. Based on our current findings, we put forward a model in which TARPs either strengthen the coupling between agonist-induced domain closure and channel opening, perhaps by promoting linker separation, or directly enhance the degree of domain closure induced by CNQX. Further structural studies are required to distinguish between these two alternatives.

References and Notes

1. T. Honore *et al.*, *Science* **241**, 701 (1988).
2. J. C. Watkins, P. Krosgaard-Larsen, T. Honore, *Trends Pharmacol. Sci.* **11**, 25 (1990).

3. M. J. Sheardown, E. O. Nielsen, A. J. Hansen, P. Jacobsen, T. Honore, *Science* **247**, 571 (1990).
4. C. J. McBain, J. V. Eaton, T. Brown, R. Dingledine, *Brain Res.* **592**, 255 (1992).
5. S. G. Brickley, M. Farrant, G. T. Swanson, S. G. Cull-Candy, *Neuropharmacology* **41**, 730 (2001).
6. G. Maccaferri, R. Dingledine, *Neuropharmacology* **43**, 523 (2002).
7. Y. Hashimoto, H. Miyakawa, Y. Kudo, M. Inoue, *Neurosci. Lett.* **358**, 33 (2004).
8. S. L. Palay, V. Chan-Palay, *Cerebellar Cortex: Cytology and Organization* (Springer, Berlin, 1974).
9. I. Tarnawa *et al.*, *Bioorg. Med. Chem. Lett.* **3**, 99 (1993).
10. K. A. Yamada, C. M. Tang, *J. Neurosci.* **13**, 3904 (1993).
11. N. A. Mitchell, M. W. Fleck, *Biophys. J.* **92**, 2392 (2007).
12. R. A. Lester, M. L. Quarum, J. D. Parker, E. Weber, C. E. Jahr, *Mol. Pharmacol.* **35**, 565 (1989).
13. R. A. Nicoll, S. Tomita, D. S. Bredt, *Science* **311**, 1253 (2006).
14. P. Osten, Y. Stern-Bach, *Curr. Opin. Neurobiol.* **16**, 275 (2006).
15. E. B. Ziff, *Neuron* **53**, 627 (2007).
16. T. A. Verdoorn, N. Burnashev, H. Monyer, P. H. Seeburg, B. Sakmann, *Science* **252**, 1715 (1991).
17. R. Jin, T. G. Banke, M. L. Mayer, S. F. Traynelis, E. Gouaux, *Nat. Neurosci.* **6**, 803 (2003).
18. K. B. Hansen, H. Yuan, S. F. Traynelis, *Curr. Opin. Neurobiol.* **17**, 281 (2007).
19. N. Armstrong, M. Mayer, E. Gouaux, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5736 (2003).
20. M. L. Mayer, N. Armstrong, *Annu. Rev. Physiol.* **66**, 161 (2004).
21. S. Tomita *et al.*, *J. Cell Biol.* **161**, 805 (2003).
22. E. S. Lein *et al.*, *Nature* **445**, 168 (2007).
23. J. R. Geiger *et al.*, *Neuron* **15**, 193 (1995).
24. M. Bertolino *et al.*, *Recept. Channels* **1**, 267 (1993).
25. K. M. Partin, D. K. Patneau, M. L. Mayer, *Mol. Pharmacol.* **46**, 129 (1994).
26. D. Catarzi, V. Colotta, F. Varano, *Med. Res. Rev.* **27**, 239 (2007).
27. M. A. Rogawski, *Epilepsy Res.* **68**, 22 (2006).
28. M. Bialer *et al.*, *Epilepsy Res.* **73**, 1 (2007).
29. A. Priel *et al.*, *J. Neurosci.* **25**, 2682 (2005).
30. S. Tomita *et al.*, *Nature* **435**, 1052 (2005).
31. D. Turetsky, E. Garringer, D. K. Patneau, *J. Neurosci.* **25**, 7438 (2005).
32. N. Armstrong, E. Gouaux, *Neuron* **28**, 165 (2000).
33. We thank P. Apostolides, J. O'Brien, R. Robbins, and L. Martin for technical assistance, and A. Tzingounis, C. Jahr, and all members of the Nicoll lab for helpful discussions. This work was supported by NIH grants to R.A.N., L'Oréal UNESCO (United Nations Educational, Scientific, and Cultural Organization) and Epilepsy Foundation predoctoral fellowships to K.M., a National Research Service Award postdoctoral fellowship (GM078754) to F.A.H., and the Center for Innovation in Membrane Protein Production (P50 GM73210).

Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5851/815/DC1
Materials and Methods
Figs. S1 to S3
Table S1
References

11 June 2007; accepted 1 October 2007
10.1126/science.1146317

Looking for a Fresh Perspective?

Join other experts and leaders
from around the globe at the
2008 AAAS Annual Meeting



**Come to Boston
14–18 February 2008**

Broaden your networks
and knowledge at the
most cross-disciplinary
conference in the world

2008 AAAS Annual Meeting

14–18 February • Boston



Special Events Include:

- Marine Sciences Seminar
- Forum for School Science
- Town Hall on Understanding Obesity and Childhood Nutrition

Highlights:

- Opening Ceremony
- 160 Symposia
- Plenary Lectures
- Topical Lectures
- Poster Sessions
- Family Science Days

Career Extras:

- *Science* Careers Fair
- Specialized Career Workshops
- Dynamic Career Resource Center

Dear Colleagues,

On behalf of the AAAS Board of Directors, it is my distinct honor to invite you to Boston for the 174th National Meeting of the American Association for the Advancement of Science.

The theme for the meeting — “Science and Technology from a Global Perspective” — emphasizes the power of science and technology as well as education to assist less-developed segments of the world society, to improve partnerships among already-developed countries, and to spur knowledge-driven transformations across a host of fields.

As you know, the AAAS Annual Meeting has become the most important gathering of the year for the growing segment of scientists and engineers who seek to explore the intersections between disciplines and to witness the broad influence of science and technology on society. **You will have the opportunity to interact with a diverse array of leading scientists, engineers, educators, students, and policy-makers.**

Attendees will have the opportunity to choose among a broad range of activities, including 160 symposia, seminars, and career development workshops as well as plenary and topical lectures by some of the world’s leading scientists and engineers. Typically the meeting includes up to 10,000 participants and hundreds of members of the national and international media.

A public-engagement event on “Understanding Obesity and Childhood Nutrition” is intended to expand the dialogue among scientists, teachers, students, policy-makers, education leaders, and the general public on the science behind childhood obesity and nutrition. It will feature a broad and exciting array of speakers with a strong focus on strategies for addressing the problem. You and your family can also enjoy **Family Science Days** — a free event open to the general public.

The following pages present the highlights of the scientific program to date. You can explore the program online, obtain updates, and develop a personal itinerary at www.aaas.org/meetings.

The Annual Meeting reflects tremendous efforts from the AAAS sections, divisions, and committees, which we gratefully acknowledge. I also extend a personal thanks to the members of the Annual Meeting Scientific Program Committee who reviewed and assembled the many excellent ideas and proposals into this outstanding meeting.

Please join us in Boston,
David Baltimore, Ph.D.

AAAS President

Robert Andrews Millikan Professor of Biology
California Institute of Technology

PHOTOS THIS PAGE COURTESY OF USDA.
PHOTOS OF RESEARCHERS OPPOSITE PAGE COURTESY OF USDA.

Plenary and Topical Speakers

Attend plenary and topical lectures given by eminent scientists and engineers.

Plenary Lectures



PRESIDENT'S ADDRESS

David Baltimore, Ph.D.

AAAS President; Robert Andrews Millikan Professor of Biology, California Institute of Technology

Baltimore is one of the world's leading biologists and a co-recipient of the 1975 Nobel Prize in Medicine for the discovery of reverse transcriptase. Since then, he has published more than 600 papers, including seminal research on the genetics of cancer, the workings of the HIV virus and AIDS vaccine candidates, and fundamental observations in molecular immunology. He was founding director of the Whitehead Institute for Biomedical Research and president of Rockefeller University and Caltech. Today he heads the Baltimore Lab at Caltech, with support from the Gates Foundation, to look for ways to genetically boost the immune system against infectious pathogens, particularly HIV. Throughout his career, Baltimore has influenced science policy. He helped set standards for recombinant DNA technology and received the 1999 National Medal of Science in part for his work on AIDS research policy. Today he is outspoken about what he sees as government efforts to distort and suppress scientific research.

President's Reception: Immediately following



Nina V. Fedoroff, Ph.D.

Special Adviser, Science and Technology, U.S. Department of State
Evan Pugh Professor of Biology and Willaman Professor of Life Sciences, Huck Institutes of the Life Sciences, Pennsylvania State University
Making the World Flat: Science and Technology in the Developing World

In August 2007, Fedoroff was named the Science and Technology Adviser to U.S. Secretary of State Condoleezza Rice. She holds an academic post as the Evan Pugh Professor of Biology and Willaman Professor of Life Sciences at Pennsylvania State University, where she is also founding director of the Huck Institutes of the Life Sciences. As a leading geneticist and molecular biologist, she has contributed to the development of modern techniques used to study and modify plants. Her book, *Mendel in the Kitchen: A Scientist's View of Genetically Modified Foods*, examines the scientific and societal issues surrounding the introduction of genetically modified crops. She received the 2006 National Medal of Science for her pioneering work on plant molecular biology and for being the first to clone and characterize maize transposons.

Plenary Lecture Panel *Global Health Challenges*

AAAS President David Baltimore, moderator

Jim Yong Kim, M.D., Ph.D.

Director, François Xavier Bagnoud Center for Health and Human Rights, Harvard School of Public Health, and Professor of Social Medicine and Medicine, Harvard Medical School (invited)

Peter Piot, M.D., Ph.D.

Executive Director, UNAIDS, and Under Secretary-General of the United Nations



JIM YONG KIM



PETER PIOT

AAAS President Baltimore will moderate a Davos-style panel discussion that explores global health challenges from three perspectives: philanthropy, world leadership, and program successes and challenges. Executive director of UNAIDS since its creation in 1995 and under secretary-general of the United Nations, Piot comes from a distinguished academic and scientific career focusing on AIDS and women's health in the developing world. Drawing on his skills as a scientist, manager, and activist, he has challenged world leaders to view AIDS in the context of social and economic development as well as security. Kim has worked to improve health in developing countries for more than 20 years and is an expert in tuberculosis. He is a founding trustee and the former executive director of Partners In Health, a not-for-profit organization that supports a range of health programs in poor communities in Haiti, Peru, Russia, Rwanda, and the United States. A panelist will be added to offer a perspective on providing support to combat diseases and to address other global health concerns.



Topical Lectures

Angela M. Belcher, Ph.D.

Germeshausen Professor of Materials Science and Engineering and Biological Engineering, Massachusetts Institute of Technology
From Nature and Back Again: Giving New Life to Materials for Energy, Electronics, and the Environment

Janet Browne, Ph.D.

Aramont Professor of the History of Science, Harvard University
2007 George Sarton Award Lecture in the History and Philosophy of Science
Commemorating Darwin: The History of Scientific Celebrations

Charles Elachi, Ph.D.

Director, NASA Jet Propulsion Laboratory
The Golden Age of Robotic Space and Earth Exploration: Challenges and Opportunities

Daniel Kahneman, Ph.D.

Eugene Higgins Professor of Psychology, Princeton University
2007 John P. McGovern Lecture in the Behavioral Sciences
Architecture of the Mind

Curtis T. McMullen, Ph.D.

Cabot Professor of Mathematics, Harvard University
The Geometry of 3-Manifolds

Per Pinstrup-Andersen, Ph.D.

H.E. Babcock Professor of Food, Nutrition, and Public Policy, Cornell University
Science and Policy Priorities for the Global Food System

Nathan D. Wolfe, Ph.D.

Professor of Epidemiology, UCLA School of Public Health
Viral Forecasting

Topical Lecture Panel

Advancing Science and Fostering Innovation Through International Cooperation: A Trans-Atlantic Perspective

AAAS President David Baltimore, moderator

Mark Fishman, M.D.

President, Novartis Institutes for BioMedical Research

Janez Potocnik, Ph.D.

Commissioner for Science and Research, European Commission



PHOTOS OF RESEARCHERS COURTESY OF USDA.

Discount Airfares to Boston

American Airlines and Delta Airlines, the official carriers for the 2008 AAAS Annual Meeting, are offering discounted airfare to and from the meeting.

For details, visit: www.aaas.org/meetings and select "Travel Information."



Explore the possibilities

Experts from 56 countries will share their work and perspectives on the breadth of science, engineering, and technology. Explore new advances, opportunities, and frontiers – all from a fresh perspective.

Climate Change and the Environment

- Biological Adaptation to a Changing Climate
- The Carbon Journey: Understanding Global Climate Effects and Advancing Solutions
- Global Interactions Between Climate Change and Microbial Activity
- Ocean Acidification and Carbon–Climate Connections: Lessons from the Geologic Past
- Ocean Iron Fertilization and Carbon Sequestration: Can the Oceans Save the Planet?
- The Other Carbon Dioxide Problem: Ocean Acidification
- Probing Arctic Regions: Linking Past Records, Present Effects, and Future Predictions
- Strange Days on Planet Ocean: New Insights on the Effects of Climate Change
- To What Extent Does Solar Variability Contribute to Climate Change?
- Transforming Our Ability To Predict Climate Change and Its Effects
- Under Thin Ice: Global Warming and Predatory Invasion of the Antarctic Seas

Communicating Science and Technology

- Communicating Science in a Religious America
- A Comparative Look at Markets, Media, and Emerging Attitudes About Nanotechnology

- A Crack in the Lab Door: The State of “Upstream Engagement” in Science
- Design of Mechanical Puzzles
- Engaging the European Public in New Science
- Global Issues: Helping the Public Understand When Scientific Information Is Valid
- Global Warming Heats Up: How the Media Covers Climate Change
- Improving Public Understanding of Engineering: From Research to Practice
- Major Transformations in Evolution: The State of the Art and Public Understanding
- New Techniques in the Evaluation and Prediction of Baseball Performance

Ecology and Resource Management

- Adaptation Options for Climate-Sensitive Ecosystems and Resources
- Designing Self-Maintaining Deltas: A Multidisciplinary Approach to Restoration
- Drylands and the City: Global Issues and Perspectives
- Finding Sustainability Without Stability: New Goals for a World in Flux
- Forum for Sustainability Science Programs
- Into the Deep: Ecology and Evolution of Deep-Sea Corals
- Planet Earth: Lessons Ignored, Lessons Learned
- Soil Protection for Sustainable Well-Being
- The Unnatural History of the Sea: New Insights and Baselines for Ocean Recovery



Visit our Web site for a listing of speakers and complete details of the meeting:

www.aaasmeeting.org

- Will Too Few Jaws Take Too Big a Bite? The Importance of Sharks to Ocean Ecosystems
- Education and the Work Force
- Aiming Higher: The How and Why of Advancing Women in Agricultural Sciences
- Defining Secondary Science Education Through Advanced Placement Redesign
- Emerging Research for Women in Science in the New Century
- Inside the Double Bind: Women of Color in STEM
- Looking Across the Ocean: Increasing Science and Engineering Women Faculty
- PISA 2006: How Well Prepared Are 15-Year-Olds for a Global Scientific Society?
- Power in Its Place: Science in Tribal Education
- Promoting the Success of Minority Graduate Students
- Science for Tomorrow's Citizens and Leaders

Food Protection and Supply

- Dolly for Dinner? Technological and Socioeconomic Perspectives of Animal Cloning
- Embracing Change: A New Vision for Management in Coastal Marine Ecosystems
- Energy, Agriculture, and People: Global Implications for Science and Policy
- Food Security and Climate Change in Africa
- Last Best Chance for Tuna: Learning from the Cod Collapse
- The Privilege To Fish

Future of Energy

- Biofuels from Forest-Based Biomass
- Biomass-to-Biofuels Conversion: Technical and Policy Perspectives
- Burn or Bury? Global Proposals for Managing Highly Radioactive Nuclear Waste
- Coal Gasification: Myths, Challenges, and Opportunities
- Food and Fuel: Biofuels, Development, and a Sustainable Bioeconomy

- Progress in Magnetic Fusion Energy Research: 50 Years of International Collaboration and Future Prospects
- Materials Synthesis Opportunities for Global Energy Needs
- Nanocatalysis for Clean Energy and Sustainability
- Nuclear Reactor Systems of the Future: Costly Illusion or Promising Breakthrough?
- A Thirst for Power: The Global Nexus of Energy and Water
- World Biofuels Production Potential in the Next Decade

Global Health: What's Next?

- Drugs in Our Corn Flakes? Our Health and the Economic Risks of "Pharma" and Industrial Crops
- Fighting the Global Obesity Epidemic: Small Steps or Big Changes?
- From Kitchen Sinks to Ocean Basins: Emerging Chemical Contaminants and Human Health
- How the Bugs Come Back and Bite Us: The Rise of Agricultural Pathogens
- Measuring Human Exposures to Hormone Disruptors: Scientific Tools for Global Health
- Overcoming the Current Challenges in HIV/AIDS
- A Perspective on Infectious Diseases: Challenges to Global Health
- Preparing for the Health Effects of Climate Change: Science and Societal Strategies
- Sustaining Human Health in a Changing Global Environment
- Understanding the Linkages Between Environmental Toxicity and Human Disease
- A Viral Time Bomb: Hepatitis C in the Developing World

Human Development and Understanding

- Advances in Language and Speech Science and Technology
- Brain Basis of Speech
- Cross-Cultural Perspectives and Mental Illness

- How Can Social Learning Move Us Toward Sustainability?
- Imagining the Future: New Perspectives from Psychology and Neuroscience
- Mathematics and the Brain
- Mind of a Toolmaker
- Moral Judgment: Evolutionary and Psychological Perspectives
- Poverty and Brain Development: Correlations, Mechanisms, and Societal Implications
- Thinking With and Without Language

Opportunities for the Developing World

- Access to Medicines: Fostering Unique Collaborations, Overcoming Challenges
- Collaboratively Developing Student Mathematical Thinking Among APEC Member Economies
- Earth Observation for Africa, with Africa
- Geospatial Science, Global Change, and Sustainability in Africa
- It Takes a Bank to Raise a Village: Financing Science and Technology for Sustainable Development
- Producing Scientists and Engineers in Developing Countries: New University Paradigms
- Progress in Human Genetics Research in Africa: Science, Technology, and Ethics





Special AAAS Membership Offer


Do you have colleagues who are not yet members of AAAS?

If they register in advance for the 2008 Annual Meeting in Boston, they will receive a one-year membership to AAAS along with all member benefits. These include a one-year subscription to the journal *Science*, online access to *Science* and all of its archives, and access to *Science Express*. International members will receive *Science Digital*.

This offer is good for advance registration only, and expires on 18 January 2008. Only nonmembers qualify.

Share the news now.

Register Now: Get special discounts on meeting and hotel registration.

Visit www.aaas.org/meetings and select  Register

- Science and Technology for Sustainable Development: The African Context
- Supporting African Scientists in the Quest for Sustainability: After Training, What?
- Our Networked World
- Blogs, Boards, and Bonding: Using Electronic Communities To Support Women in Science
- Building Science Capacity with Linked Observation Systems: Seismological Perspectives
- Challenges in Collecting and Interpreting Data in Humanitarian Emergencies
- Enhancing Science Globally Through High-Performance Computing and Simulation
- From Space to Village: Promoting Sustainable Development Using Satellite Observations
- The Global Dimension of Research Infrastructures
- Information, Computing, and Communications: Keys to Sustainable Global Development
- Managing and Preserving Scientific Data: Emerging Perspectives on a Global Basis
- Power of the Internet To Facilitate Science Education and Networking: The Supercourse
- Seeing Science
- Virtual Observatories and Research Collaboratories: Network-Enabled Science
- Science and Engineering on the Horizon
- 50 Years of the Space Age: Looking Back, Looking Forward
- Earth Observations from Space: 50 Years of Accomplishments
- Global Diffusion of Nanotechnology: Lessons from China, Italy, and the United States
- Grand Challenges and Opportunities for Engineering in the 21st Century
- High-Powered Lasers: Fusion Ignition and Concomitant Scientific Opportunities

Airport Transportation

For information about transportation from the airport, see www.aaas.org/meetings and select "Travel Information."

- Mars Rovers: The Exploration of Mars
- New Diamond Age: From Optics and Electronics to Spintronics and Nucleonics
- Nuclear Physics: New Answers, New Questions About the Visible Universe
- Quantum Information Theory
- What Is a Planet?
- Worldwide Hunt To Solve the Mystery of Gamma-Ray Bursts

Science, Technology, and Public Policy

Advocacy in Science: Opportunities, Limits, Responsibilities, and Risks

- Art and Connoisseurship: New Scientific Techniques Conserve Art and Architecture
- Optimal Laboratory Animal Care and Use: The Road to International Guidelines
- Strengthening Federal Science Through the 2009 Presidential Transition
- Toward a New Climate Economics: Can We Afford the Future?
- Turning the Tide? Current Climate and Energy Technology Policy Negotiations
- We the People: Funding Science Through Direct Democracy
- Where Does "Think Globally, Act Locally" Come From?
- Whose Scientific Judgment?

The Scientific Enterprise

- Ethical Issues in Scientific Publishing
- From Global to Local: Impact of Field Research in Biological Anthropology
- Global Research Competitiveness in Times of Social Transition
- Is It Possible To Predict the Future of Science?
- Research and Technology at the Crossroads of the Debate on Biopiracy
- Sustainability of Open Access: Does Increasing Global Access Come with Hidden Costs?
- Translation of Fundamental Cancer Biology: Toward Clinical Innovation — Singapore Model
- Unlocking the World's Science: Increasing Access, Adding Visibility, and Aiding Authors

Strategies for International Scientific Cooperation

- Are There Diverse Paths to Progress in Global Science?
- Changing Models of Research in Higher Education: International Perspectives
- Engaging the World Through Science: Science Partnerships in U.S. Foreign Relations
- English-Only Science in a Multilingual World: Costs, Benefits, and Options
- Global Knowledge and Information Commons for Sustainability Science and Innovation
- Global Partnerships for Sustainability Science
- Global Science in the Modern World: Perspectives on the Quantum Universe
- Humanity at the Nexus: Academic Partnerships and Entrepreneurship in Global Health
- Interdisciplinary Research and Integrated Policy-Making for Sustainable Development
- Large-Scale International Collaborations and the Future of Physics
- Universities Without Walls: Endeavors in Global Interinstitutional Education

Technology for a Healthy Future

- Air Pollution and Atherosclerosis
- Bisphosphonate Therapy and Oral Problems, a Two-Edged Sword?
- Crops for Health: Improving the Health-Promoting Properties of Food
- The Father and the Fetus: Revisited
- Health Economic Evaluations of Medical Technologies: Is the Cost Worth the Cure?
- High-Tech, Low-Cost Medicine: A New Paradigm for Global Health
- Modeling the Dynamics of the Drug-Resistant Killers of the 21st Century
- Nanotechnology and Health: What Are the Benefits and Risks?
- A New Generation of Studies To Unravel the Genetic Components of Cardiovascular Diseases
- Progress in Cancer Prevention
- The Science Behind Consumer Product Safety Testing: New Directions
- Systems Biology: Hype or Hope in Drug Design?



World Security and Stability

- Atomic Detectives: Nuclear Forensics and Combating Illicit Trafficking
- Biometrics in Border Management: Grand Challenges for Security, Identity, and Privacy
- Global Ecologies of Danger: Living Through Extreme Times
- Nuclear Forensics and Global Nuclear Deterrence
- Radiation Detectors for Global Security: The Need for Science-Driven Discovery
- Technologies for Open Source Intelligence: Staying Ahead of the Game
- Terrorist Threats to the United States



PHOTO OF RESEARCHER COURTESY OF USDA.

Students: Take the freeway to Boston Attend the meeting for free ... volunteer as a session aide.

Check out what you'll get if you volunteer:

- A free subscription to *Science* Online
- Free access to the Career Resource Center
- Free meeting registration

For more information:
www.aaasmeeting.org/student



Discover the Power of Networks

Start your own global conversations

Engage in powerful networking opportunities. Hear and discuss the latest advances in scientific research and technology.

Understanding Obesity and Childhood Nutrition: A Special AAAS Public Engagement Event

Teachers, school health professionals, parents, students, scientists, and the public are cordially invited to take part in a free town hall-style event on understanding the science behind obesity and childhood nutrition. The town hall will explore key questions, including:

- › Is it the fat, the carbohydrates, or the calories?
- › What does science tell us about children's nutritional needs worldwide?
- › Is there enough time in the school day for exercise?
- › How can nutrition instruction fit into the K-12 science curriculum?
- › What roles can communities, schools, culture, and people play in addressing the problem of childhood obesity?

Program updates and more information will be available at www.aaas.org/obesity.

2008 Forum for School Science Programs that Create a New Science Professional: The Ph.D. as Public Educator

What do Georges Charpak, Russell Hulse, Leon Lederman, and Carl Wieman have in common aside from physics and the Nobel Prize?

All are research scientists who are engaged in K-16 science education: Charpak designed La Main à la Pâte, a hands-on teaching method for primary grade students. Hulse served as a visiting professor of physics and science and mathematics education at University of Texas at Dallas. Lederman created the Illinois Science and Math Academy (one of many things). Wieman moved to the University of British Columbia to focus on improving undergraduate physics education.

What motivates science professionals to get involved in science education and to engage in collaborations with an individual teacher, a school, a school system, or a science organization in the community? The 2008 Forum for School Science will:

- › Describe selected collaborations and partnerships between universities and school systems,
- › Provide steps for developing and sustaining the partnerships,
- › Describe the value added for the institutions and individuals involved, and
- › Offer data and analyses from projects and outside evaluators on the effectiveness of the program and the impact on graduate students, teachers, the universities, and the school systems.

The Forum will highlight programs that build partnerships between institutions of higher education and school systems to promote improvements in science, technology, engineering, and mathematics education, and engages a new generation of scientists to communicate their research to the public.

Managing Threats to Marine Ecosystems

Organized by: Larry Crowder, Duke University, Beaufort, N.C.

This two-day seminar focuses on threats to marine ecosystems, first describing the location and magnitude of those threats based on new spatially-explicit analyses. New sensing technologies and new analytical tools map the effects of human activities on marine resources as well as collateral damage to habitats and non-target species.

Because of these new place-based approaches, management efforts can be focused where they are likely to make a difference. Valuation of ecosystem services allows scientists and managers to examine the utility of various management approaches, including marine reserves, to protect ecosystem functioning and resilience.

Attendees will hear about a variety of new discoveries and novel methodologies that will fuel the development of marine ecosystem-based management.

Discount Hotel Rates

AAAS has negotiated special rates for AAAS Meeting attendees. For rates and details, visit www.aaas.org/meetings and click on "Hotel Reservations." Find out why it pays to book a room through AAAS.



Broaden Your Reach

If your organization is looking to make a strong impression on a large scientific, technological, and engineering audience, consider exhibiting or sponsoring at the 2008 AAAS Annual Meeting in Boston. Join a list of world-class sponsors, supporters, and exhibitors.

AAAS wishes to thank the following sponsors and supporters:

Presenting Sponsor



SUBARU

*P&G*beauty



L'ORÉAL



CALL YOUR VIDEO PROVIDER
FOR MORE INFORMATION



THE KAVLI PRIZE



Johnson & Johnson
PHARMACEUTICAL RESEARCH
& DEVELOPMENT, L.L.C.

SUPPORTERS: Argonne National Laboratory and Yomega

In addition generous funding for AAAS Awards is provided by Johnson & Johnson Pharmaceutical Research & Development, L.L.C. and Affymetrix.

EXHIBITORS (as of 1 October)

Academia Book Exhibits
American Mathematical Society
Baden-Württemberg International
Basic Books
Brookhaven National Laboratory
Council for International Exchange of Scholars
ESO – European Organization for Astronomical Research in the Southern Hemisphere
EurekAlert!
European Commission, DG Research
European Molecular Biology Laboratory
EuroScience
Hokkaido University
Institute of Physics Publishing
Irish Universities Association
John Templeton Foundation
Marine Conservation Alliance
Massachusetts Institute of Technology

NASA Earth Systems Science — Data and Services
National Aeronautics and Space Administration
National Center for Science Education
National Center for Supercomputing Applications
National Institute of Standards and Technology
National Oceanic and Atmospheric Administration
National Science Digital Library
National Science and Engineering Research Council of Canada
National Science Foundation
New Scientist/Reed Business Information
Oak Ridge Associated Universities
Oak Ridge National Laboratory
Penguin Group (USA)

Proceedings of the National Academy of Sciences
Prometheus Books
RIKEN, Japanese Research Institute
Science
The Science Channel
Subaru of America Inc.
U.S. Civilian Research and Development Foundation
U.S. Department of Science, Genomes to Life Project
U.S. Department of Science, ITER Project
U.S. Department of Science, Retina Project
U.S. Department of Science, Science.gov
U.S. Department of Science, Workforce Group
U.S. Environmental Protection Agency
University Corporation for Atmospheric Research

For more information:

Contact: Jill Perla
AAAS Meetings Department
Direct Dial: 202-326-6736
E-mail: jperla@aaas.org



ADVANCING SCIENCE. SERVING SOCIETY